UIUCDCS-R-74-677

HARDWARE DESIGN OF A REAL-TIME MUSICAL SYSTEM

BY

SERGIO FRANCO

October, 1974

Department of Computer Science
University of Illinois
Urbana, Illinois 61801

This work was submitted in partial fulfillment of the requirements for the degree of Doctor of Philsophy in Computer Science, at the University of Illinois.

HARDWARE DESIGN OF A REAL-TIME MUSICAL SYSTEM

Sergio Franco, Ph.D.
Department of Computer Science
University of Illinois at Urbana-Champaign, 1974

ABSTRACT

This thesis describes the hardware design of a hybrid system for the composition and performance of electronic music in real-time. While analog circuitry is primarily employed in the generation and processing of sounds, digital circuitry is devoted to the exercise of control, under the immediate supervision of the composer/performer. Such a digital vs. analog partition, together with a proper choice of the man vs. machine interface, is intended to satisfy the much emphasized musical need for the immediate, real-time interaction between the composer/performer and his instrument.

ACKNOWLEDGMENTS

The author expresses his sincere gratitude to Prof. Michael Faiman, his thesis advisor--and friend--for his indispensable support, advice and suggestions throughout the development of this thesis.

The author is also pleased to credit Dr. J. Divilbiss with initiating him into the fascinating field of electronic music.

This thesis has emerged from a two-year collaboration with composer Salvatore Martirano and with the other members of the project. To his friend Sal, the author expresses his warmest thanks for those endless, heated and constructive discussions that have made this thesis possible.

The diligent typing and the beautiful drawings of this thesis have been made by Ms. Evelyn Huxhold and under the supervision of Mr. Mark Goebel, respectively, both of whom the author thanks for their precious collaboration.

This work is dedicated to Diana.

TABLE OF CONTENTS

I. INTRODUCTION

1.1 Importance of Real-Time in Electronic Music.

With the advent of electronic music, the roles of the composer and the performer have been placed in a new perspective. Traditionally, musicians have tended to specialize in either one role or the other, or have tended to play only one of the roles at a time. This has happened mainly because of the considerable difficulty encountered when composing and performing are carried on simultaneously. When a musician improvises at an instrument, he is essentially composing music in real-time.[1] This means he is thinking and evaluating very quickly a series of possibilities out of which he makes appropriate choices, which in turn lead to the execution of the various mechanical motions necessary to produce the music he wants. Because of the difficulties encountered in doing all these things at once, musicians have tended to specialize either in the inception/evaluation process (composition) or in the execution process (performance), and have developed the notation of musical scores to communicate among themselves. Thus, the composer need not be principally concerned with the quick thinking and decision making of the improvisational process, but can pursue his inspection of musical possibilities on his own time scale. What is even more important, to appreciate the effects of his choices and changes, he does not need to try them out on actual instruments because he can rely on his aural imagery, a faculty he has developed through experience either by playing instruments or by listening to others playing them, or both.

In developing his aural imagery, the musician is certainly helped a lot by the fact that he deals with a limited set of instruments whose characteristics extend over known and predictable ranges. In electronic music the

situation is quite different. Because of the new kinds of sounds that electronic instruments are capable of synthesizing, it is very difficult for the musician to develop the kind of musical imagery that may suffice to assist him in composing on his own time scale and away from his instruments. If the composer is to come up with something musically meaningful at all, it is absolutely essential that he actually hear the effects of his choices, his trials, and his alterations as he makes them, so that he can directly evaluate them in the context of the whole composition. In other words, the concept of direct feedback has come to play a dominant role in the electronic music composing process.

Another important change brought about by electronic music involves the relationship between performer and instrument, and is due to the much greater detail in which the performer is required to control the various musical parameters. While in conventional instruments such parameters as timbres, attacks, decays, etc. are, to a large extent, fixed, built-in features, in electronic music they are left to the discretion of the performer, who must therefore specify and control them directly. If this feature allows on the one hand much more freedom of choice and experimentation, on the other it imposes a more demanding control burden upon the performer, and it certainly renders the communication among musicians more complex than with conventional instruments, where the notation of the musical score is usually adequate.

As a result of the above discussion, it should be clear that in electronic music the distinction between composer and performer ceases to exist, partly because the composer needs to try out his musical ideas on the instrument by himself, and partly because of difficulties of communication among different people. Furthermore, having accepted the notion of direct feedback as an indispensable ingredient of the improvisational process, it is of paramount

importance that the composer/performer be allowed to interact with his instrument on the time scale of the music he is improvising, that is, in real-time.

1.2 Electronic Music Instruments.

The instruments available today to the electronic musician fall into one of two categories: Analog music synthesizers and digital computers.

Analog synthesizers are self-contained systems consisting of a number of basic building blocks which can be interconnected in a variety of configurations to generate and modify sounds.* The basic sound-generating device is the oscillator, which produces a predetermined number of waveforms, usually the sinusoidal, pulse, saw-tooth and triangular waves. Starting with these basic sounds, new waveforms can be obtained by means of mixers, filters, non-linear signal processors, and other sound modifying devices. Another basic sound modifier is the attenuator, which is used to control the dynamic characteristics of the amplitude or loudness of sounds.

In order to allow musical parameter control, these circuits are designed so that their characteristics can be altered externally by the operator, a task normally accomplished by means of manual switches, potentiometers, piano-like keyboards, patch cords, etc.** Recent advances in circuit technology and design allow several circuit parameters to be electronically controllable by

---

* During the '60's, the analog synthesizer scene was mostly dominated by the instruments of R. A. Moog, operating on the East coast, and D. Buchla, on the West coast. In recent years, encouraged by the enormous progress of technology, other firms have joined in this venture, so that the market today offers more than half a dozen lines of synthesizers, spanning a wide range of quality, complexity, and price. Besides MOOG and BUCHLA, other synthesizers widely known today are ARP and PUTNEY.

** Because of the strong competition, most synthesizer manufacturers neither publicize their designs nor provide circuit schematics with the purchase of their equipment. To the author's knowledge, the only comprehensive exposition of analog synthesis concepts and design techniques available to date is offered in the papers by R. A. Moog (see references 2, 3).

voltage signals.[2,3] Hence, certain control responsibilities can be removed from the operator and built into the system by having some circuits control parameters of the others directly. In spite of this, a considerable amount of manual operation is still required of the performer, mostly in the form of knob turning and dial setting.[4,5]

With lack of sophistication and flexibility in the control of the musical parameters, the amount of influence the musician can exert upon his instrument in real-time is rather limited. This broadening of the gap between amount of control required and conditions for its implementation results in a further aggravation of the basic dilemma of the improvisational process already discussed.

Much more powerful than analog synthesizers, both in terms of parameter richness and control sophistication, are musical systems based on digital computers. Here, under the direction of appropriate programs, a computer is instructed to simulate a series of instrument building blocks, such as digital oscillators, filters, adders, multipliers, attack generators, etc., from which the musician assembles the instruments for his musical piece.[6] The computer calculates and outputs the desired sound waves in binary sampled-data form. Hence, D/A converters must be used to convert the samples into electrical pulses, and appropriate low-pass filters are needed to smooth the pulses and produce the continuously varying electrical voltages suitable for driving a system of loudspeakers. With the power of modern digital computers it is not difficult to envision the synthesis of virtually any sound that could come from a loudspeaker, which is a fairly general sound source available today. Therefore, the sound qualities obtainable by the computer method span a much richer range than in the case of analog synthesizers. Having at his disposal the resources of a whole computer, the musician can also make use of much

more sophisticated algorithms for the control of musical parameters and the organization of compositional elements.

However, when it comes to real-time performance capabilities, computers, like analog synthesizers, suffer from severe limitations, albeit of a different nature.[7] Apart from time delays determined by the particular mode of operation of the computing center--batch processing, off-line D/A conversion, etc.--there are certain limitations which are inherent to the computer itself. Since the computer is actually controlling the air-pressure that is coming directly from the loudspeaker, it must reconstruct every minute portion of the acoustic waveform of the final sound, a computational task that involves not only a lot of resources, but also a lot of time. While with analog synthesizers the generation of waveshapes in the audio range is a rather trivial matter that presents no real-time problems, with computers this same task becomes the most demanding in terms of resources and computation time. Thus, even if the musician had at his disposal a dedicated, general-purpose computer and was allowed to communicate with it in an interactive mode, there is very little he could do in real-time, just because the computer is unable to keep up with all the computations that need to be carried out on that time scale.*

If we could somehow relieve the computer from the burden of computing the waveshape samples, and were to use it only to control the perceptual characteristics of musical sounds like timbres, pitches, durations, loudness crescendi, tempi, etc., then it would certainly be able to keep up with the necessary computations in real-time.[7] This is so because the rates at which these psychological events occur or change are much lower than the rates at which the air-pressure vibrates.

---

* For a more detailed description of the equipment and techniques available to the electronic musician today, see ref. 8.

To conclude, it seems that digital systems are particularly suited to musical parameter control in real-time, while their real-time capabilities for sound synthesis are rather limited. Analog synthesizers, on the other hand, while incorporating only rudimentary facilities for parameter control, can cope very easily with the generation and modification of acoustic waveshapes in real-time, although the ranges in sound quality are not as rich as in digital synthesis. If one could improve the sophistication of analog sound generators, it would seem that a satisfactory musical system with real-time performance capabilities can be arrived at by hybrid techniques, where analog circuitry is used primarily for sound generation, and digital circuitry is devoted to parameter control. This, indeed, is the idea at the basis of the system to be described next.

1.3 General Concept of New Instrument.

If the assurance of a real-time interaction between composer/performer and instrument is to be the primary objective, two conditions must be met: both operator and machine must be capable of carrying out their respective tasks in real-time.

Real-time capabilities for the machine are achieved by resorting to a hybrid configuration, as mentioned in the previous section. Thus, the generation of sounds and, in general, the performance of those tasks that, because of the great amount of computation associated with them, take too long with digital computers, are assigned to analog circuitry. The exercise of control, on the other hand, is carried out more efficiently with digital circuitry. To make parameter control possible, the characteristics of analog circuits are externally programmable. The programming is done directly by the digital portion of the system, under the control supervision of the composer/performer.

Knowing the limitations of analog synthesis with respect to sound quality, one may argue that a hybrid system, as far as this musical feature is concerned, is going to be limited at least by the extent to which the analog portion is. This is indeed true. However, with the help of state-of-the-art technology and careful design, the sophistication of sound generating and modifying circuitry can be improved considerably. A constant effort in this direction has been made throughout the development of the system, from the design of the pitch generators to the design of the timbre control, loudness control, sound distribution system, and others. Due to the presence of analog circuitry another problem arises, namely, parameter stability. The human ear is very sensitive to effects that result from subtle relationships and proportions among certain musical quantities: tuning, beats, dissonances, simultaneity of events, etc. These psychoacoustical factors make a severe demand upon analog circuit performance, a problem usually not so serious in all-digital systems. In order to alleviate these shortcomings as much as possible, particular emphasis has been placed on the maximization of accuracy and stability, especially where the demand for these features is more critical.

Once the conditions for real-time operation of the machine are satisfied, we must ensure that the human operator, too, be able to keep up with his tasks on the improvisational time-scale. For this to be possible, he must be relieved of the burden of too much control detail, which requires an amount of activity that is usually incommensurate with human efficiency. Human operation in real-time can be best achieved if the relationship of the performer/composer to the system is like that of a conductor to his orchestra, rather than that of a player to his instrument. Instead of taking care of every elemental musical constituent, as one would normally do with electronic music instruments and, to a lesser extent, with traditional instruments, the improviser should mainly limit himself to the activity of steering, guiding, and

influencing the evolution of a system that is already, so to speak, capable of playing by itself.

If the performer is to expect such a degree of cooperation from the system, the latter must be provided with the capability of controlling its own parameters automatically. The dynamic control of compositional elements--be it automatic or under the immediate responsibility of the operator--requires the generation of a continuous stream of binary information. Information for atuomatic control is generated within the system by means of binary sequence generators based on feedback shift-registers and read-write memories.

Besides being capable of automatic control, the system must also be receptive to some form of external control from the operator. Since our goal is that of relieving the latter from too much control detail, the compositional elements that are best suited to the direct control of the improviser, both in terms of their musical significance and the associated data rate, are those affecting music at the macrostructural level. The control of microstructural elements, such as individual fluctuations in the perceptual characteristics of sounds, involves a data rate that is generally too high in relation to human capabilities in real-time, and is therefore handled more efficiently by the system itself. The operator, however, can still affect microstructural elements indirectly through the usage of the control power he has at the macrostructural level.

The above concepts can be better illustrated by examining the stream of control information involved in the production of a typical musical piece. As already mentioned, the dynamic control of perceptual characteristics of sounds is handled with sequences of binary control words. Typically, perceptual characteristics fluctuate at rates of the order of 1 to 100 times per second, which clearly shows that the operator can hardly keep up with the

associated data rate, especially if several perceptual parameters are to be controlled simultaneously. Hence, the generation of this type of control sequences is delegated to the system, which can certainly keep up with the required rate without difficulties.

Examination of the stream of information over a longer period of time, as, for instance, over the span of a complete musical gesture, is likely to reveal the presence of another class of fluctuations which affect the information at the level of entire sequences. Sequence rests and activations, fluctuations of sequence lengths and rates, changes in the ranges spanned by sequence data words are examples of this class of macroscopic, or second-order fluctuations which affect music at a higher level of structural complexity. The manipulation of macrostructural elements itself presupposes the need for a corresponding class of control sequences to carry out the task, thereby leading to the notion of sequences controlling other sequences. Owing to the multilevel structural complexity of music, the chain of sequences controlling sequences of the level below can be expanded further. However, we need not elaborate on this topic here, as our goal is directed solely to pointing out the real-time implications of such a distinction between different classes of sequences. In this respect, what characterizes macroscopic fluctuations is the lower rate at which they take place, which may be typically of the order of one event every 1 to 100 seconds. This is of the same order of magnitude of the rate at which a human can comfortably perform control actions in real-time. Therefore, it seems rather appropriate to give the performer direct access to the control of this class of fluctuations. It may be noted, in passing, that the notion of sequences controlling other sequences allows a more uniform treatment of control and a higher degree of standardization in circuit design and implementation.

The generation of sequences for the control of macrostructural fluctuations need not be a prerogative of the performer exclusively. Indeed, in the course of an improvisation he may be willing to seek the automatic co-operation of the system in carrying out even this type of task. To ensure maximum flexibility, it is desirable that the balance between man and machine be allowed to vary continuously and under the control of the operator himself. This feature brings about a distinction between automatic and manual information which is instrumental in the design concept of the system. The first refers to the information produced within the system by sequence generators, as already mentioned. The second refers to the information that the operator feeds into the system through the input interface. Whether at any moment a given compositional element is to be controlled automatically or manually is decided by the performer with the aid of a simple information-steering device.

1.4 System Organization.

Figure 1 shows an overall system block diagram. Both in analog and digital terms the instrument can be partitioned into four distinct layers or subsystems. These subsystems, however, are not totally independent of each other, since they are allowed to share common resources, to exchange information, and to be put in hierarchical relationships with respect to each other. This system partition is intended to reflect the musical notion of four distinct orchestras which can play together and interact in a variety of ways and at different levels. As a result, the performer is given more flexibility in achieving contrasting musical situations than with a monolithic, homogeneous system.

A common resource that is always allocated to one orchestra at a time is the control panel, which constitutes the interface between performer and instrument. The panel consists of an array of touch-sensitive switches with

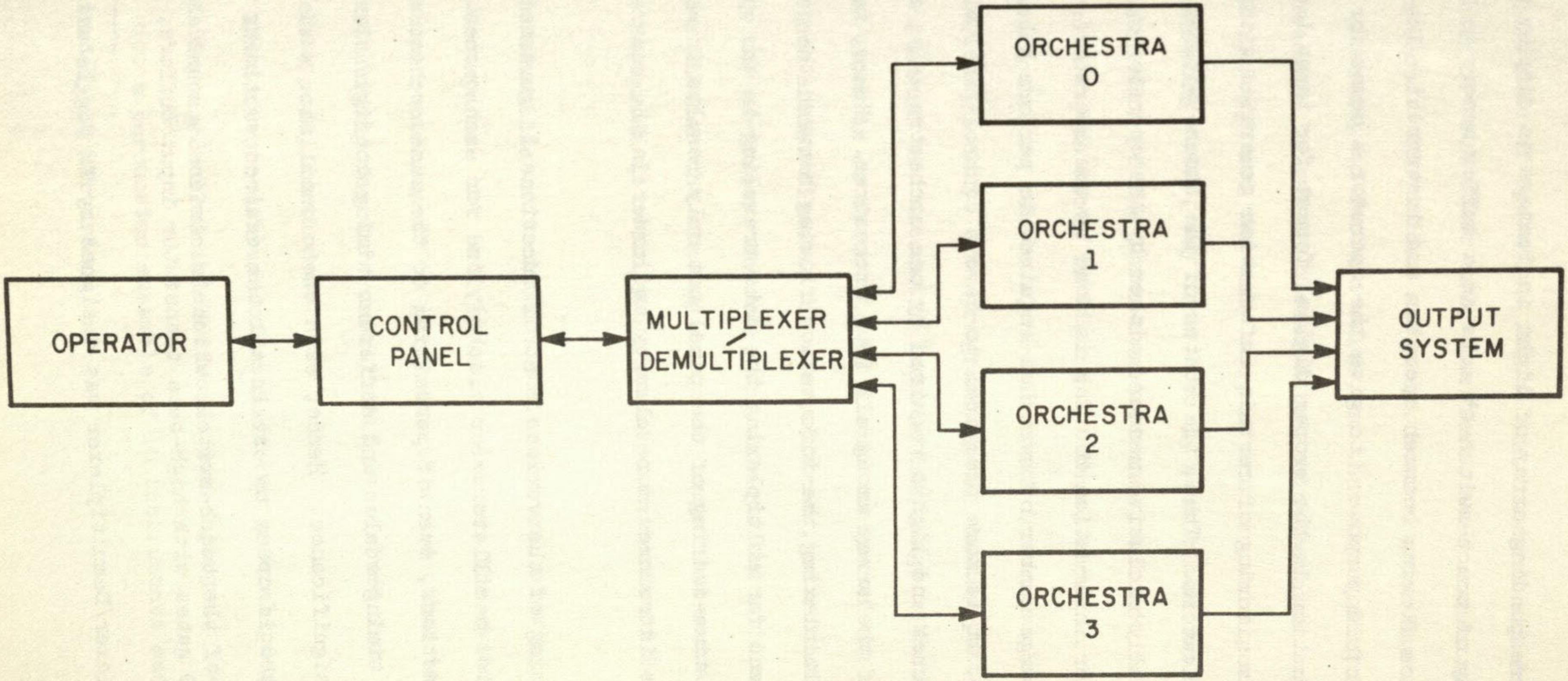


Figure 1. System Block Diagram. Orchestras interact among each other. For clarity, inter-orchestra arrows have been omitted.

memory, and a corresponding array of light indicators to display the switch states. The array of touch-switches* acts as a buffer memory to handle the digital information flowing between operator and instrument. The control panel serves a twofold purpose: to give the operator a means for setting up his control information in the proper digital format for input into the system, and to allow the monitoring of certain information generated within the system, via the light indicators. Thus, the states of the touch-switches can be controlled in a mutually exclusive manner, either by the operator or by the system, and the proper information routing is done by the operator.

Due to the large number of switches required to perform all the desired control functions, it is more practical to have a control panel with just the array of switches and lights required by one orchestra only, and to time-share the usage of the array among all four orchestras. Hence, facilities are provided for demultiplexing the information flowing between the panel and the four subsystems, and for multiplexing the data traveling in the opposite direction.** The time-sharing of the panel not only results in reduced cost and size, but also introduces more clarity and order in the control activity of the performer.

The basic timing of the various control functions is governed by four digitally-controlled oscillators (or clocks), one for each orchestra. Different control functions, even if pertaining to the same orchestra, in general require different timing scales and different timing configurations, depending on their musical significance. Hence, each basic oscillator must be properly scaled and processed in order to obtain all the desired subtiming signals.

* The design of the touch-switch, which is centered around a pair of cross-coupled NAND gates with high-beta transistor input buffers, is due to Dr. J. Divilbiss.

** The Multiplexer/Demultiplexer was designed by T. Noggle and R. Borovec.

The simplest mode of operation is the one in which all four basic clocks run independently of each other and, therefore, each orchestra plays according to its own tempo. Provisions are made, however, for the slaving of clocks to each other according to certain rules, so that the timing systems of different orchestras can be put in hierarchical relationships. A meaningful slaving rule,* although not the only one, is that in which, with the clocks numbered in ascending order, say from 0 to 3, each clock can be slaved to any one or combination of the lower-numbered ones, but not to any of the higher-numbered ones, as shown in Figure 2. Thus, clock 2, for instance, can run independently, or can be slaved to clock 1, or to clock 0, or to clocks 1 and 0, but not to clock 3. The programming of slaving patterns offers a good example of manual vs. automatic information usage. Whether at any moment a given clock is or is not to be slaved to one of the other eligible clocks is decided either directly by the performer, or automatically by the system. It is the performer, however, who selects the mode of operation, that is, who decides whether the slaving is to be controlled manually or automatically. This mechanism is readily implemented in terms of 2-to-1 data selectors.

Since the automatic control of the hierarchical slaving is a decisional problem that encompasses the individual orchestras, the system is provided with a supervisor control unit to carry out this and other tasks of similar type. The basic timing for this unit itself is derived from a combination of the four clocks already mentioned.

Each of the four orchestras consists of two general-purpose instruments, or voices, as they are called in electronic music jargon. Two of the orchestras include also a percussion ensemble of 16 instruments each.

* Conceived by S. Martirano.

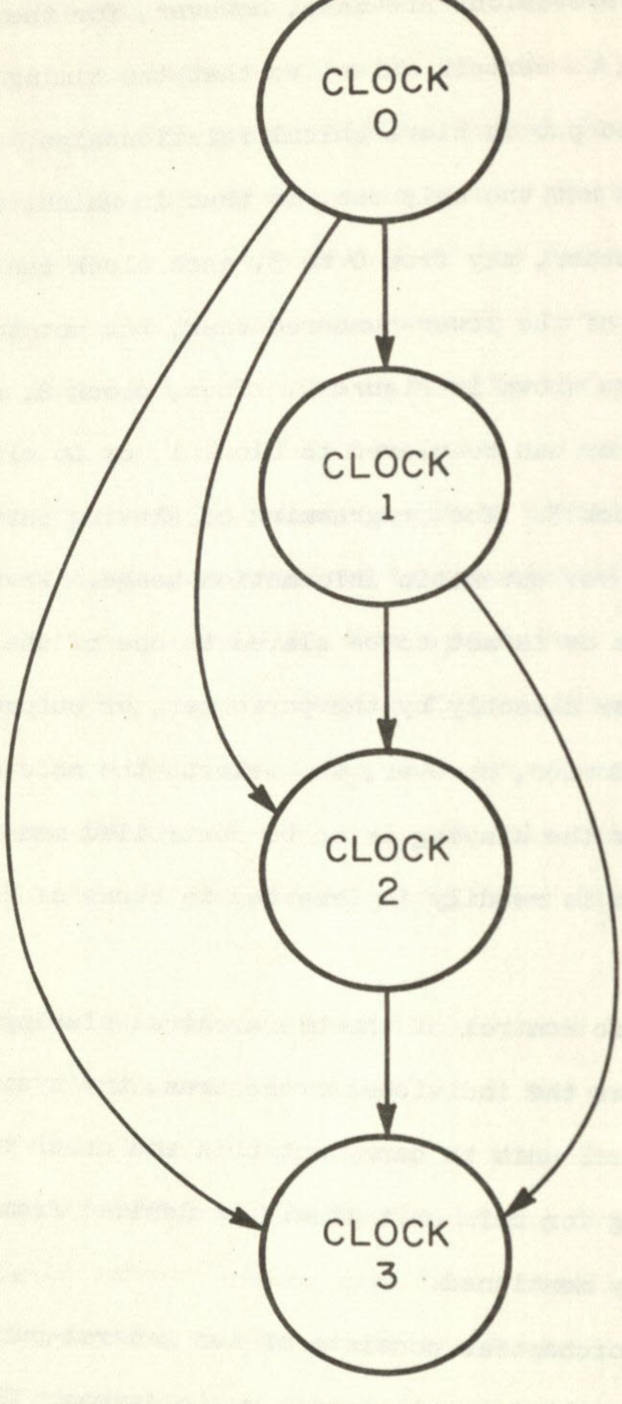


Figure 2. Clock Hierarchy.

A voice module is shown in block diagram form in Figure 3. It consists of a Frequency Synthesis device for the manipulation of pitches; a dual Digital Waveshape Generator for the programming of timbres; a Mixer/Modulator and a Programmable Filter for the dynamic control of tonal characteristics; an Attenuator/Locator for the instrumental as well as spatial characterization of sounds.

The eight voice modules are coupled with each other in a variety of ways, so that, under command from the control system--be it manual or automatic--the electronic signals associated with one voice device can be utilized to modulate parameters of others. For instance, one waveshape may be used to modulate the amplitude, the pitch, or the timbre of another. This form of analog parameter control is a feature that the system offers in addition to the digital form of control already mentioned. Indeed, since the whole system is hybrid, it comes as no surprise that control can be exercised in digital as well as in analog form whenever this is possible and musically effective.

Although percussive sounds can be readily synthesized by properly programming the parameters of voice modules, they are actually generated separately by means of simpler, special-purpose circuitry. This solution is intended to reduce the overload of the voice modules, so that they can be more efficiently allocated to the synthesis of sounds of higher complexity.

Almost all circuits are dc coupled, a feature that allows a certain amount of circuit standardization. Thus, the basic clocks are nothing else but waveshape generators as in the voice modules, the only difference being in the values of timing capacitors, which are properly scaled in order to obtain the desired parameter ranges. In this case the Frequency Synthesis device serves the purpose of controlling durations and associated timing functions, while the sub-audio waveshapes here obtained are utilized to

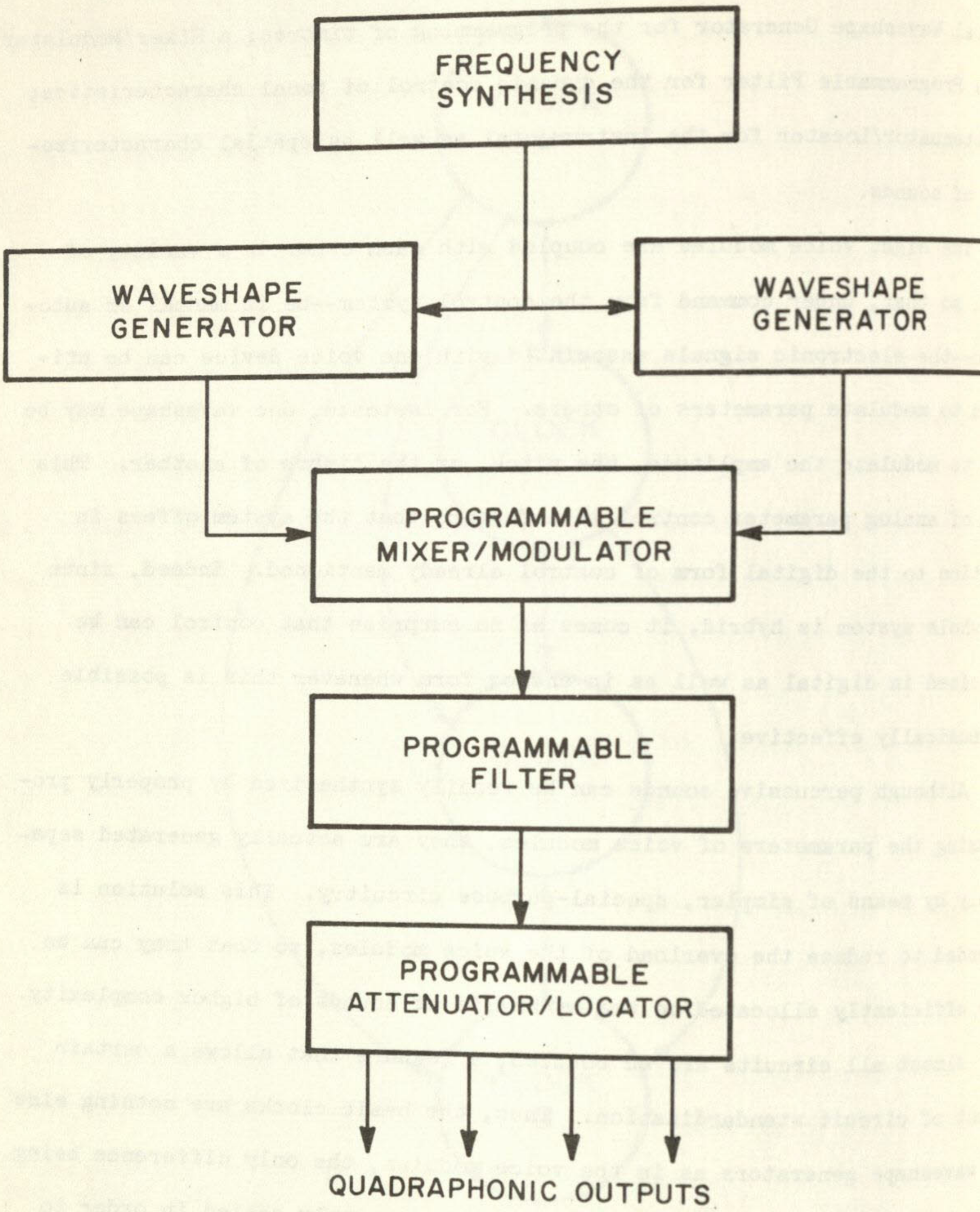


Figure 3. Voice Module Block Diagram.

control slow-changing musical effects and parameters like crescendi, glissandi, transpositions, timbres, portamento, etc. (Figure 4).

Another resource shared by all four orchestras is the output system which carries the sounds created within the instrument to the outside world. For performances confined to a limited space or for studio recordings, the instrument operates in a quadraphonic mode. Each orchestra controls the directionality of its sounds dynamically by continuously programming their distribution among the four channels. This task is carried out by means of programmable joystick-like devices, which constitute part of the Attenuator/Locator appearing in the block diagram of Figure 3. Facilities are also provided for the control of the depth of sounds, besides their directionality, through the use of programmable artificial reverberation.

For performances taking place in a concert hall, the spatial resolution of the quadraphonic system can be made more subtle and sophisticated by resorting to a larger number of loudspeakers properly distributed among the audience. In this case, the output system is switched to an alternative mode of operation which involves the use of 24 loudspeakers having 4 input channels each--one for each orchestra--for a total of 96 channels, as shown in Figure 5. The presence or absence of sound on any of these channels is controlled by a corresponding audio gate. Thus, every orchestra can control the routing of its own sounds to any combination of the 24 loudspeakers according to patterns that can be programmed directly by the performer on the control panel. Such spatial effects as sound travel and spatial dialogues can be easily implemented with this type of sound distribution.

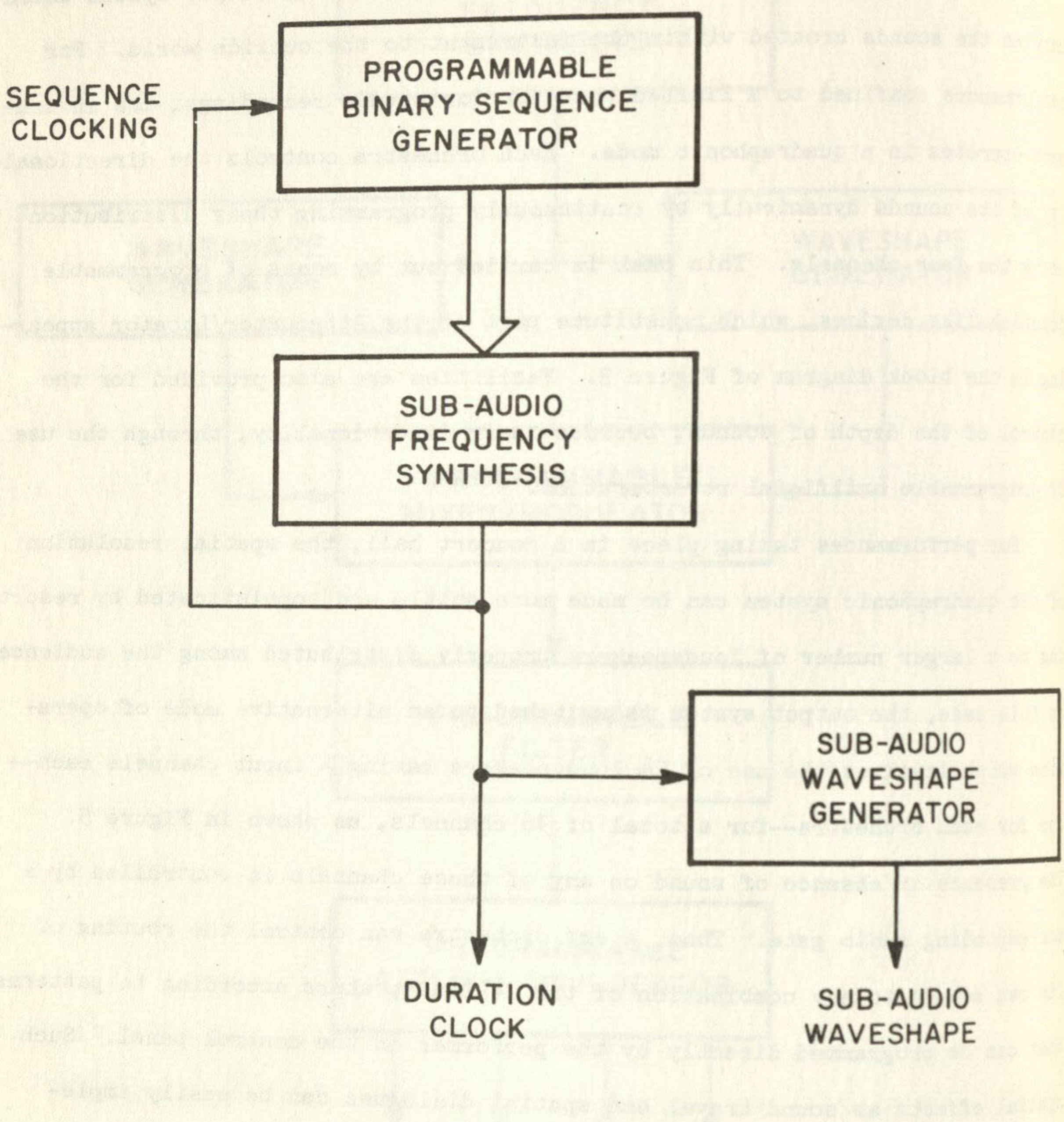


Figure 4. Duration Clock Generation for One Orchestra.

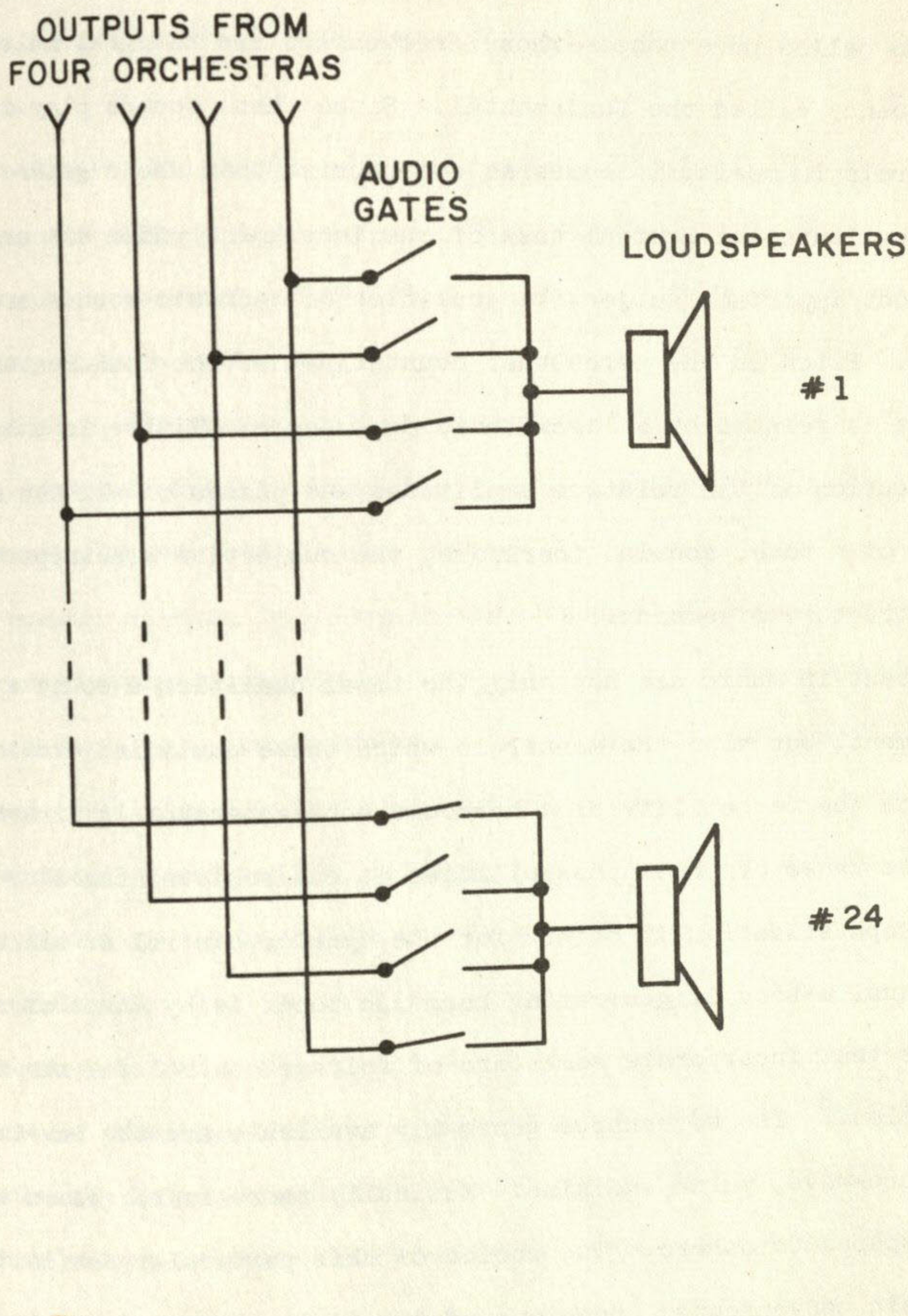


Figure 5. 96-Channel Output System.

II. TONE SYNTHESIS

2.1 Harmonic Tone Generation.

Harmonic tones are periodic sounds consisting of a mixture of sinusoidal waves--also called pure tones--whose frequencies are integral multiples of a basic frequency called the fundamental. Since these sounds play a rather important role in music, it comes as no surprise that their generation and control constitute the central task of the instrument under discussion.

The most important subjective qualities of harmonic sounds are pitch and timbre. Pitch is the perceptual counterpart of the fundamental frequency, to which it is related by a logarithmic dependence. Timbre is the result of the distribution of the relative amplitudes and phases of all the significant components of a tone, and is, therefore, the subjective counterpart of the spectral structure of sounds.

Important in music are not only the tonal qualities a sound exhibits at a given moment, but also the manner in which these qualities evolve with time. Hence the versatility of a harmonic tone generator is directly related to the range of tonal possibilities it can achieve, together with the amount of sophistication it offers for the dynamic control of the tones.

The usual method of generating harmonic tones is by means of relaxation oscillators that incorporate some form of voltage control for the frequency of oscillation.[8] The waveshapes generally available are the saw-tooth, triangle, square-wave, pulse and sine. Musically there is no reason to prefer these waveshapes to others. The choice of this particular set is dictated by electronic convenience: once one of the first two waveforms is available--and this is usually the case with relaxation oscillators--all others can be easily obtained from it by means of relatively simple circuits. Additional

waveforms can be generated by mixing waves of the basic set together, or by altering their spectral structure through the use of formant filters or non-linear signal processors. Although expanding the initial set of basic wave-shapes, these techniques afford an amount of tonal variety and a sophistication of dynamic tonal control that are still inadequate to span a broad musical scope.

In view of this limitation, an alternative approach to the synthesis of harmonic tones was adopted here, namely, the generation of waves in sampled form. This method, as shown in the block diagram of Figure 6, is centered around a binary counter and a random-access memory. The waveshape is digitized and stored in the memory, with each memory word holding the value of one waveshape sample. Memory addressing is provided by the binary counter, whose modulus equals the total number of words in memory. A train of pulses feeding the counter results in a sequential, repetitive scan of all memory words which are thereby displayed at the memory sense outputs and converted into analog voltages by a D/A converter. Thus, what is ultimately obtained is a staircase version of the waveshape stored in memory. In order to eliminate the unwanted discontinuities introduced by the finite sampling resolution, the staircase is subsequently smoothed by a low-pass filter.

It is immediately seen that pitch can be controlled merely by programming the frequency of the pulse train feeding the counter. To maintain a uniform amount of smoothing that is independent of pitch, the position of the filter cut-off frequency relative to the fundamental frequency of the waveshape must be kept constant over the whole frequency range. Thus the filter is also programmable, and its tuning is controlled by the same mechanism that programs the frequency of count.

II. TONE SYNTHESIS

2.1 Harmonic Tone Generation.

Harmonic tones are periodic sounds consisting of a mixture of sinusoidal waves--also called pure tones--whose frequencies are integral multiples of a basic frequency called the fundamental. Since these sounds play a rather important role in music, it comes as no surprise that their generation and control constitute the central task of the instrument under discussion.

The most important subjective qualities of harmonic sounds are pitch and timbre. Pitch is the perceptual counterpart of the fundamental frequency, to which it is related by a logarithmic dependence. Timbre is the result of the distribution of the relative amplitudes and phases of all the significant components of a tone, and is, therefore, the subjective counterpart of the spectral structure of sounds.

Important in music are not only the tonal qualities a sound exhibits at a given moment, but also the manner in which these qualities evolve with time. Hence the versatility of a harmonic tone generator is directly related to the range of tonal possibilities it can achieve, together with the amount of sophistication it offers for the dynamic control of the tones.

The usual method of generating harmonic tones is by means of relaxation oscillators that incorporate some form of voltage control for the frequency of oscillation.[8] The waveshapes generally available are the saw-tooth, triangle, square-wave, pulse and sine. Musically there is no reason to prefer these waveshapes to others. The choice of this particular set is dictated by electronic convenience: once one of the first two waveforms is available--and this is usually the case with relaxation oscillators--all others can be easily obtained from it by means of relatively simple circuits. Additional

waveforms can be generated by mixing waves of the basic set together, or by altering their spectral structure through the use of formant filters or non-linear signal processors. Although expanding the initial set of basic waveshapes, these techniques afford an amount of tonal variety and a sophistication of dynamic tonal control that are still inadequate to span a broad musical scope.

In view of this limitation, an alternative approach to the synthesis of harmonic tones was adopted here, namely, the generation of waves in sampled form. This method, as shown in the block diagram of Figure 6, is centered around a binary counter and a random-access memory. The waveshape is digitized and stored in the memory, with each memory word holding the value of one waveshape sample. Memory addressing is provided by the binary counter, whose modulus equals the total number of words in memory. A train of pulses feeding the counter results in a sequential, repetitive scan of all memory words which are thereby displayed at the memory sense outputs and converted into analog voltages by a D/A converter. Thus, what is ultimately obtained is a staircase version of the waveshape stored in memory. In order to eliminate the unwanted discontinuities introduced by the finite sampling resolution, the staircase is subsequently smoothed by a low-pass filter.

It is immediately seen that pitch can be controlled merely by programming the frequency of the pulse train feeding the counter. To maintain a uniform amount of smoothing that is independent of pitch, the position of the filter cut-off frequency relative to the fundamental frequency of the waveshape must be kept constant over the whole frequency range. Thus the filter is also programmable, and its tuning is controlled by the same mechanism that programs the frequency of count.

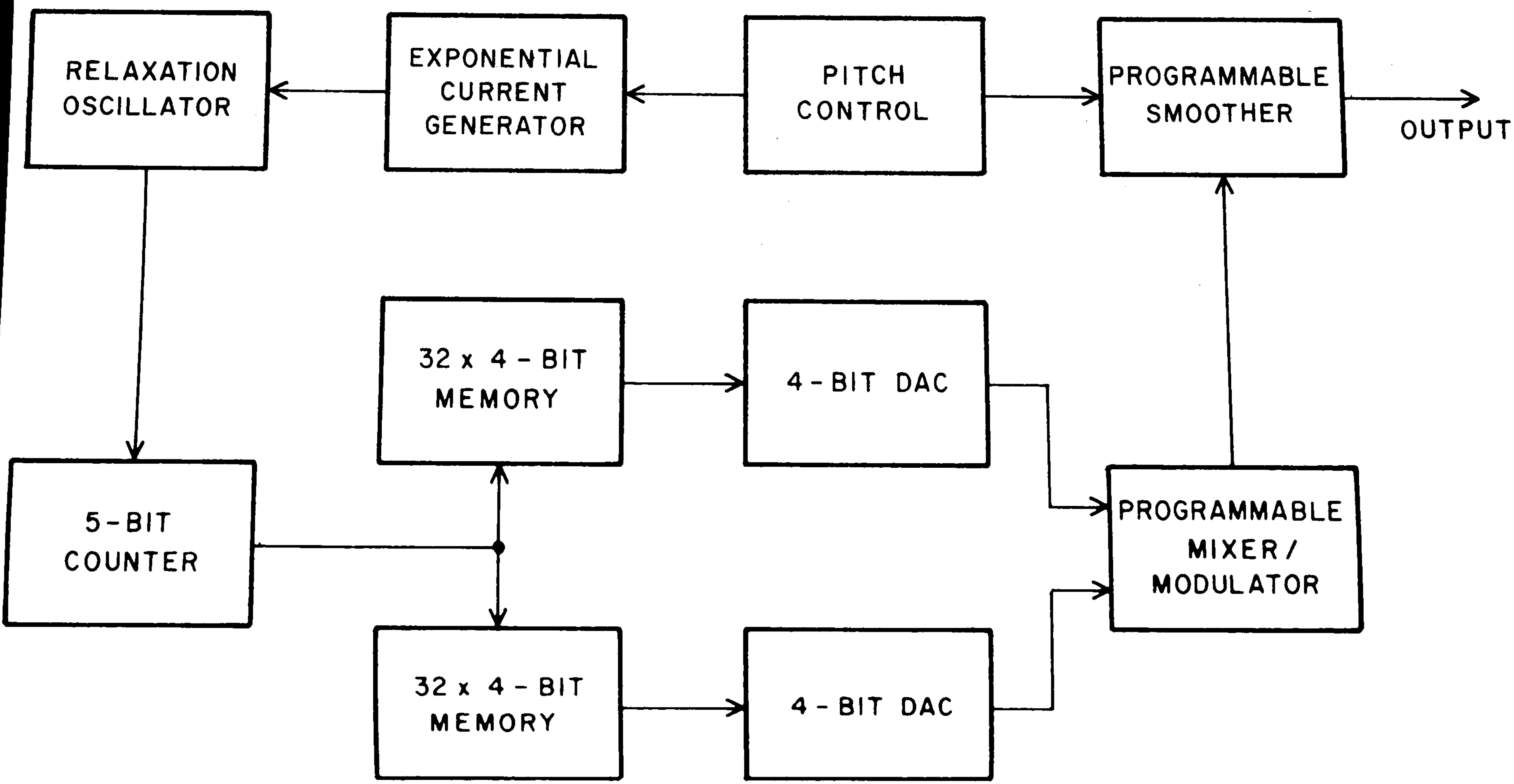


Figure 6. Harmonic Tone Generator.

The shape of the wave emerging from the filter is determined by the information residing in memory. Being provided with read/write capabilities, the memory can be reprogrammed at will to accommodate new waveshapes. Hence, the same basic circuitry lends itself rather easily to the generation of a rich family of harmonic tones.

Usually a memory-write operation results in an abrupt change from one waveshape to another. From the viewpoint of tonal control, however, it is generally desirable that waveshapes undergo smooth transitions, although abrupt changes may be occasionally acceptable as an option.

If the initial and final waveshapes are available simultaneously, a smooth transition between the two can be easily implemented in analog fashion by means of a continuously programmable mixer. The simultaneous generation of two waveshapes is achieved by using two separate memories instead of one. To ensure that the waves emerging from the corresponding D/A converters have identical pitch, albeit different shapes, both memories are addressed by the same counter. The two waveforms are then fed to the analog mixer where they are combined in programmable, complementary percentages. Thus, the application of a linear sweep to the mixer programming input causes the emerging waveshape to undergo a smooth glide from one of the incoming waveshapes to the other. At the beginning or at the end of a sweep, therefore, the contribution to the mixer output comes entirely from one of the two memories, while the contribution from the other is under complete attenuation by the mixer itself. By allowing either memory to be updated only when its contribution is totally attenuated, the discontinuities associated with memory-write operations are always kept below audibility. Consequently, the listener perceives waveshape changes only in the form of smooth transitions.

This technique of dynamic spectral control can be put to good use, among others, for enhancing the expressiveness of synthesized sounds. It has long been recognized that electronic sounds generally lack the warmth and life that characterize natural sounds. To a large extent this difference stems from the fact that the spectral structure of a natural sound changes considerably in the course of its duration, however brief it may be.[9] While computer sound generation does not lack effective approaches to the problem of spectral control,[10] the techniques employed in analog synthesis are somewhat artificial and unsatisfactory. It is seen that the hybrid approach to sound synthesis here adopted allows for flexible control over the dynamics of audio spectra.

The versatility of the waveshape generator can be further expanded by allowing for alternative modes of memory read/write control. Thus, facilities are provided for overriding the constraints illustrated above and for updating a memory even when the corresponding output is not necessarily under complete attenuation by the mixer. This option is particularly useful in view of the fact that, when in the write-mode, the information present at the data inputs of the memories here used is transferred directly to the sense outputs. If the input information is derived, for instance, from a pseudorandom binary generator, then one obtains colored noise, and the amount of coloring can be easily controlled with the programmable, low-pass filter available at the output. By forcing only one of the memories into the write mode while the other is allowed to proceed in the usual way, the mixer output consists of pitched noise, and the amount of pitching can be programmed to arbitrary values with the aid of the mixer itself.

These options can be properly exploited for the creation of a variety of special effects, as will be discussed in more detail below. Also below,

it will be shown how the mixer, besides serving the purposes illustrated above, can be used to implement various forms of amplitude modulation, thereby expanding the class of available sounds to include anharmonic ones.

2.2 Pitch Control.

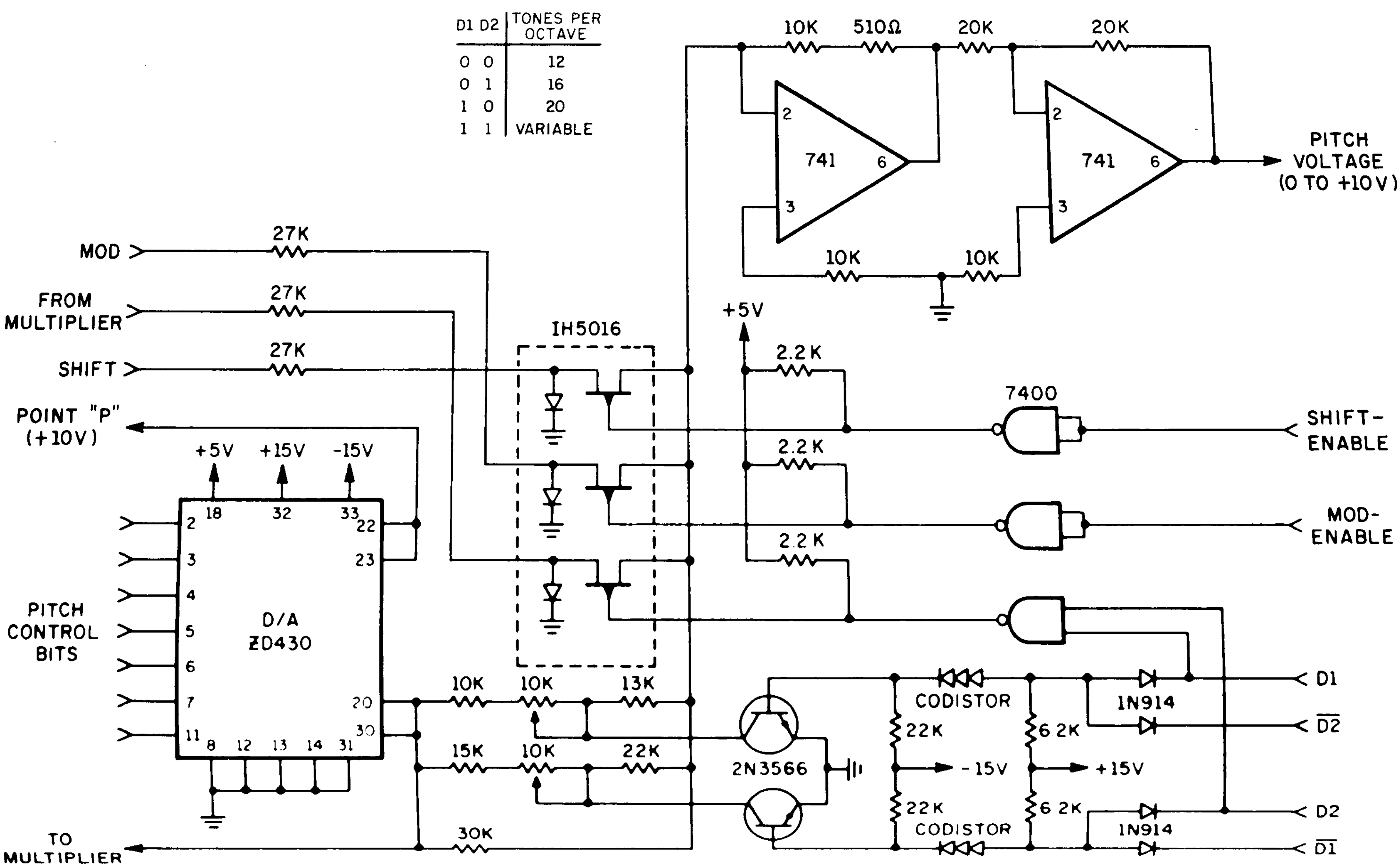
The purpose of the pitch control circuit, as shown in Figure 7, is the generation of a voltage signal in the range 0 to +10V, which will ultimately be mapped into a pitch in the audio range.

Seven bits of information yield 128 different pitches which, in the basic mode of operation, implement an equitempered scale of 12 tones per octave. Besides this basic mode, the system can operate also with 16, 20, or a variable number of equitempered tones per octave. The variable number can be any integer or non-integer value between 16 and 20. All of these options could be readily implemented by means of a multiplying D/A converter. However, because of the still prohibitive prices of these devices at the time of design--1972--a D/A converter of fixed full-scale was used instead, and the proper scales are obtained with the help of a digitally-programmable amplifier in conjunction with an external, variable-transconductance, analog multiplier. Independently of the equitempered mode in which the system is operating, the circuit also incorporates inputs for uniform pitch transposition and for frequency modulation.

The reference voltage to set the scale of the D/A converter as well as the scales of the circuits to be described in the following two section are derived from the temperature-compensated 10V voltage source internal to the D/A converter itself.

2.3 Exponential Current Generation.

Owing to the fact that the frequency intervals of interest in music form a geometric rather than an arithmetic progression, the linear domain



D1 D2	TONES PER OCTAVE
0 0	12
0 1	16
1 0	20
1 1	VARIABLE

Figure 7. Pitch Control.

of the pitch control-voltage just discussed must be mapped into an exponential domain. This conversion is accomplished by exploiting the relationship between collector current I_C and voltage drop v_{BE} across the base-emitter junction of a silicon transitor, which is expressed by the well known formula[11,12]

$$I_C = \alpha_F \, I_{ES} \, [\exp\,(qv_{BE}/kT) - 1] \; . \qquad (1)$$

Here α_F is the forward, short-circuit current gain, common base; I_{ES} is the emitter saturation current; q the electron charge; k Boltzmann's constant; T the absolute temperature.

The dynamic range of interest for pitches is about 3 decades, or 10 octaves. For reasons that will become apparent later, the corresponding range for exponential current I_C has been chosen to lie between 1 μA and 1 mA. For a high-beta transistor operating over this range, Equation (1) can be replaced, at the price of negligible error, by the truly exponential relation

$$I_C = I_{ES} \exp\,(qv_{BE}/kT) \; . \qquad (2)$$

Since the human ear is extremely sensitive to pitches out of tune, the utmost care must be exercised to ensure a stable and accurate exponential generation. Thus, compensation for temperature variations of I_{ES} is achieved by resorting to a pair of transistors lying on the same substrate and, therefore, satisfying the conditions for close matching and thermal coupling.[12] As shown in Figure 8, the exponential conversion proper is done by transistor Q3. Transistor Q4, being driven at constant current by regulator A1, develops a base-to-emitter voltage drop which is also constant, apart from temperature variations tracking analogous variations in Q3. Thus Q4 provides the appropriate low-impedance, temperature-tracking emitter drive for exponential generator Q3.

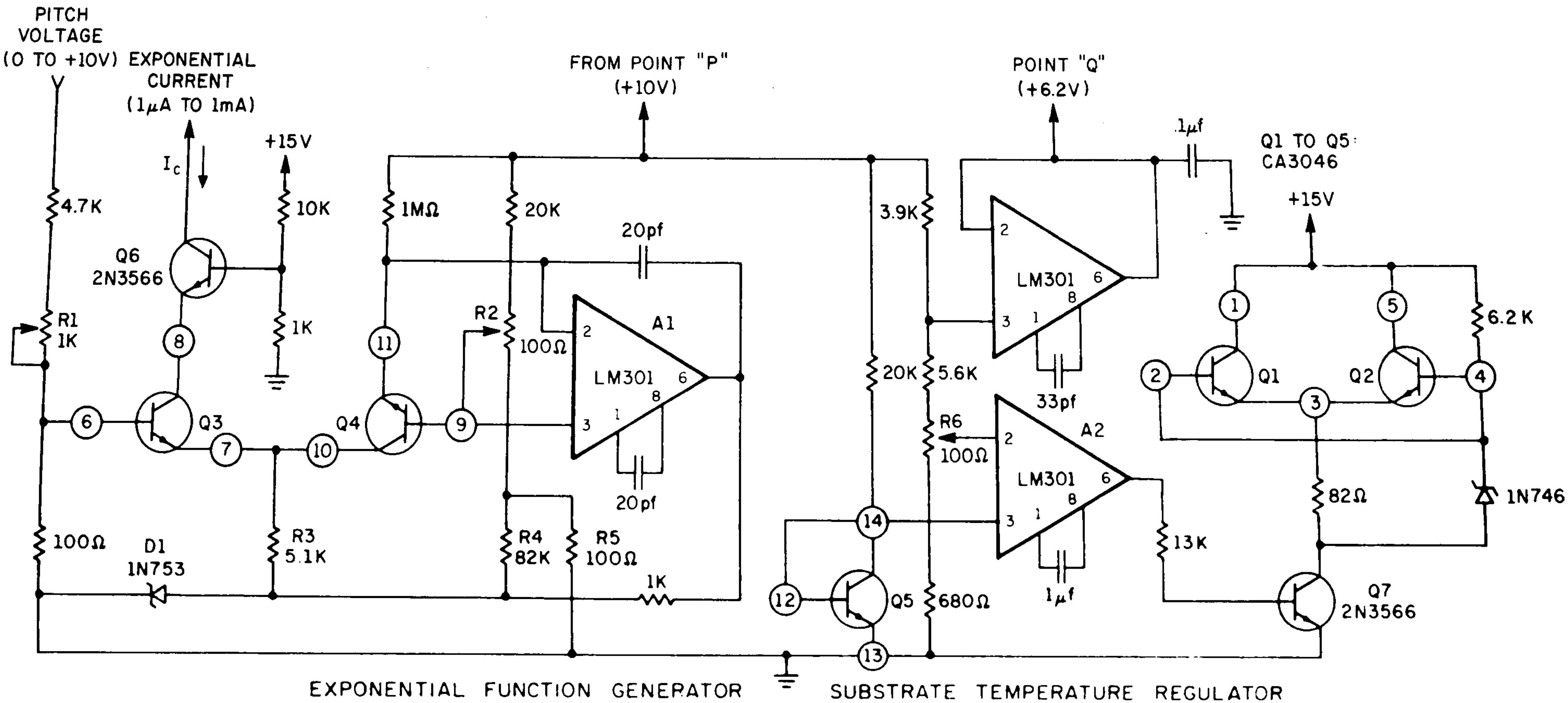


Figure 8. Temperature-Compensated Exponential Current Generator.

Compensation for temperature variations of I_{ES} does not eliminate thermal fluctuations entirely, however. As Equation (2) shows, I_C still depends on T through the exponential term. To minimize this last source of instability, the temperature of the transistors' substrate is maintained constant by another regulator circuit. The RCA 3046 transistor array, besides the pair assigned to the exponential function generation, includes three additional transistors. One of these, Q5, is connected as a diode and is used by external regulator A2 to sense the IC substrate temperature. The remaining two, Q1 and Q2, merely dissipate power in order to keep the substrate temperature constant at a value set by external trimpot R6. This value is chosen so that the quiescent power dissipated by Q1 and Q2 into the substrate, at normal ambient temperature, is half its maximum. Transistors Q1 and Q2 draw about 30 mA of current and operate in a constant-current mode in order to avoid large load changes for the positive power supply.

Having minimized thermal instability, there remains one last source of error to be taken care of. This is introduced by the bulk resistance r_E of the emitter region of Q3, which causes the junction voltage v_{BE} to differ from the voltage V_{BE} actually measured between base and emitter terminals, by the amount $r_E I_C$. If this fact is taken into account, then Equation (2) becomes[11]

$$I_C = I_{ES} \exp\left[q(V_{BE} - r_E I_C)/kT\right] ,$$

which shows that the relationship between I_C and V_{BE} is not truly exponential. As r_E is only of the order of 10 ohms, the bulk voltage $r_E I_C$ is inconsequential in the lower part of the I_C range. Not so in the upper portion of the range, where, for instance, $r_E I_C$ is no longer negligible in comparison with a V_{BE} excursion of about 18 mV, which is the amount typically required to effect a change of one octave in I_C.

This source of inaccuracy can be compensated for by applying a corrective voltage of magnitude $r_E I_C$ either to the base of Q3, in the positive direction, or to the base of Q4, in the negative direction. As a negative voltage proportional to I_C is already being developed across emitter resistor R3, the desired correction is achieved merely by scaling such voltage down and by feeding it to the base of Q4, a task accomplished by resistor divider R4 and R5.

Trimmers R1 and R2 set the width and the base of the I_C range respectively. Zener diode D1 clamps the upper value of I_C in the case of input overdrive, and acts therefore as a filter for pitches beyond approximately 20 kHz. To avoid accuracy degradation, the exponential generator is buffered to the subsequent stage by high-beta, common-base transistor Q6.

2.4 Relaxation Oscillator.

The task of this stage is to convert the exponential current into a proportional frequency which in turn is used to drive the counter addressing the memory. In order for the final waveshape to lie in the audio range, the counter must count N times as fast as the intended audio frequency, where N is the modulus of the counter. As the counter used is five bits long, this means the exponential frequency must lie approximately between 640 Hz and 640 kHz, a range which presents unusual problems if accuracy is to be our primary objective.

Current-to-frequency conversion is achieved by means of the relaxation oscillator shown in Figure 9. Assuming for a moment that R1 = 0 and capacitor C is discharged, current I_C charges C and causes the voltage at the lower side of C to decrease from the initial value of 6.2 V. This voltage is buffered by the high input-impedance, unity-gain, FET amplifier to comparator A1 which compares it with a threshold voltage of 1.9 V as set by resistor

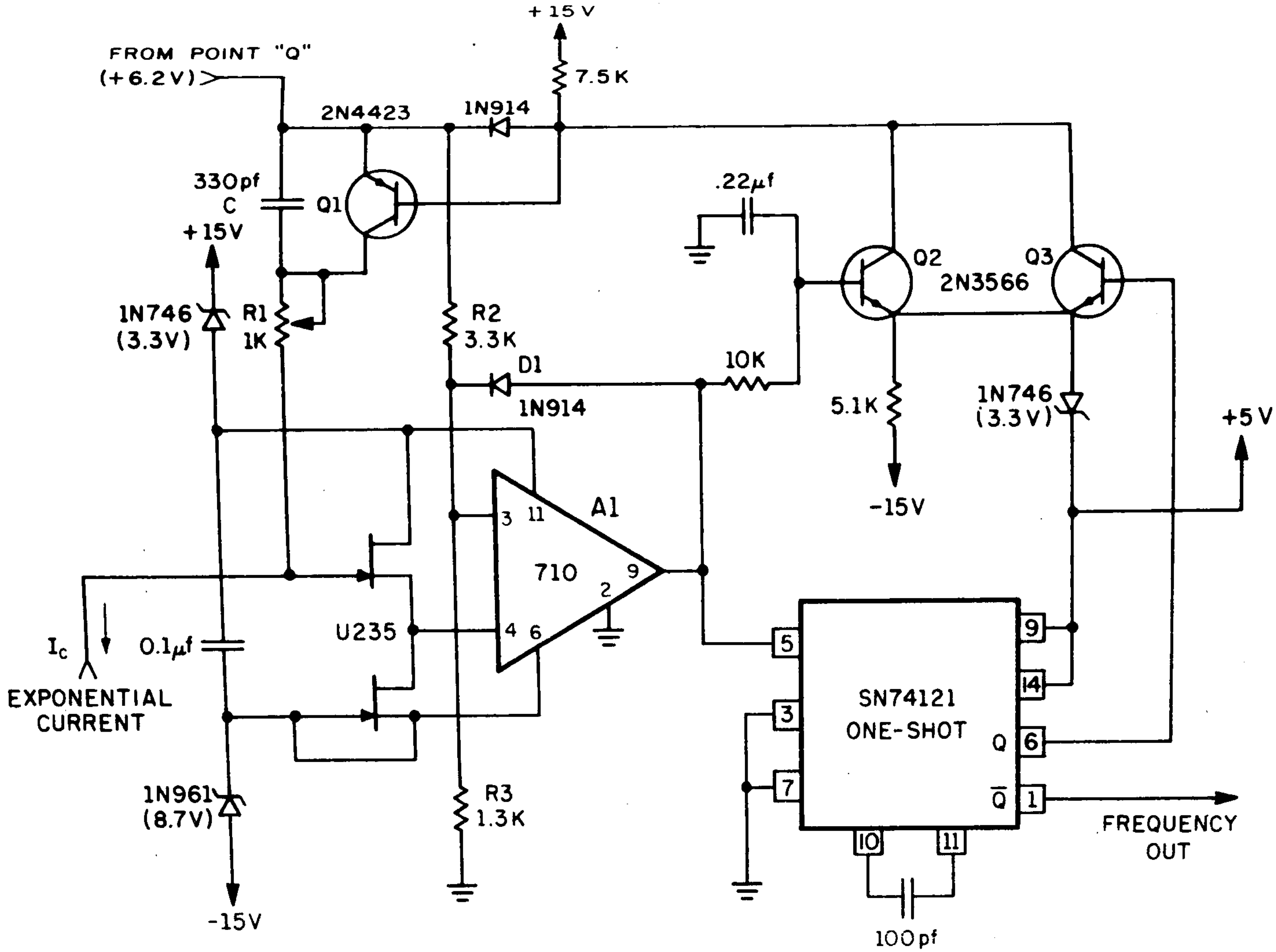


Figure 9. Relaxation Oscillator.

divider R2 and R3. As the capacitor voltage hits the threshold, A1 triggers the one-shot which in turn causes switch Q1 to close and discharge C. The delay-constant of the one-shot is about 150 ns, which is long enough to allow for the complete discharge of C by Q1. After the one-shot time-out, switch Q1 is opened again and the cycle is repeated.

The introduction of the one-shot between A1 and Q1 serves the purpose of establishing a reliable delay between the triggering of A1 and the initiation of a new cycle. Thus, fluctuations of the time taken by Q1 to discharge C are completely masked by the stable delay-constant of the one-shot.

The overall delay τ occurring between the triggering of A1 and the beginning of a new cycle, although inconsequential at low frequencies, is no longer negligible in the upper frequency range, where periods are of the order of microseconds. This fact causes the relationship between output frequency and input current I_C to be no longer linear. The error introduced by the presence of τ can be compensated for by causing comparator A1 to trigger a time τ prematurely, so that the overall duration of a cycle can be restored to its proper value. Such compensating action is performed by resistor R1. Being driven by current I_C, R1 develops voltage drop $R1 \cdot I_C$ in series with the voltage across capacitor C. If V_C is the capacitor voltage that causes A1 to trigger, then

$$6.2\ \text{V} - V_C - R1 \cdot I_C = 1.9\ \text{V} .$$

The time T_C it takes to charge C to voltage V_C at constant current I_C is

$$T_C = C \cdot V_C / I_C .$$

Elimination of V_C yields

$$T_C = 4.3 \cdot C / I_C - R1 \cdot C .$$

The duration T of an entire cycle is then

$$T = T_C + \tau = 4.3 \cdot C/I_C - R1 \cdot C + \tau \ .$$

By choosing R1 so that $R1 = \tau/C$, the frequency $f = 1/T$ becomes

$$f = I_C/(4.3 \cdot C),$$

i.e., the frequency is linearly related to current I_C over the whole range.

It is worth noticing that, to a first-order approximation, the error contributed to f by τ is of the same type as the error contributed by the exponential generator bulk-resistance r_E, as a series expansion of the corresponding formulas will quickly reveal. Thus, rather than having a separate trimpot for the bulk-resistance compensation, a fixed resistor divider was used there, and the final overall trimming for both compensations is done by means of R1 alone.

In order to maintain the fluctuations of stray-capacitances negligible in comparison with the value of capacitor C, the latter should be chosen as large as possible. The larger the capacitor, however, the larger must be current I_C, in order to obtain the prescribed frequency range, and this is undesirable in view of the error caused by bulk-resistance r_E. A compromise was achieved by choosing C = 330 pF and I_C in the range 1 μA to 1 mA.

To improve noise and jitter immunity, comparator A1 drives the one-shot through its Schmitt-trigger input, and A1 itself is connected as a Schmitt-trigger by virtue of diode D1.

Transistor Q3 serves the purpose of interfacing the one-shot to switch Q1, while transistor Q2 ensures proper start when power is turned on.

The overall accuracy of the oscillator frequency vs. pitch-voltage exponential relationship was measured to be better than ± .2% over the prescribed three decades of range. This figure is about 1/30th of a semitone in the conventional scale.

2.5 Digital Waveshape Generation.

Because of the discretization introduced by the sampling technique, the control of tonal qualities can be exercised only over a limited bandwidth. According to the well known sampling theorem, the order of the highest harmonic that can still be represented with the sampling process equals half the number of samples contained within one waveshape period. Thus, if the harmonic generator is to span a wide tonal scope, the number of sampling intervals should be as large as possible. A high sampling density, however, requires a commensurably high frequency for the sequential scan of memory as well as a large memory size to accommodate the samples. As it has already been pointed out in connection with the relaxation oscillator, the accurate generation of an exponential frequency over a range of three decades becomes more and more difficult as the range is shifted in the direction of higher frequencies. A compromise has been achieved with the choice of a time resolution of 32 samples per period and an amplitude resolution of 16 levels per sample. As the tonal bandwidth is determined solely by the number of samples, the amplitude resolution need not be as high as the time resolution. It should be noted, however, that the presence of the mixer results in an effective increase of the amplitude resolution because it allows for the programming of arbitrary ratios between corresponding samples of the incoming waveshapes.

With the above specifications, the realization of a digital waveshape generator requires four SN7489 IC bipolar memories, as shown in Figure 10. The open-collector memory outputs are fed to simple diode-resistor D/A converters which generate the low-level analog signals--0 to approximately +32 mV--suitable for driving the programmable mixer. The inputs to the mixer are provided in balanced form to eliminate the d.c. offset of about 16 mV

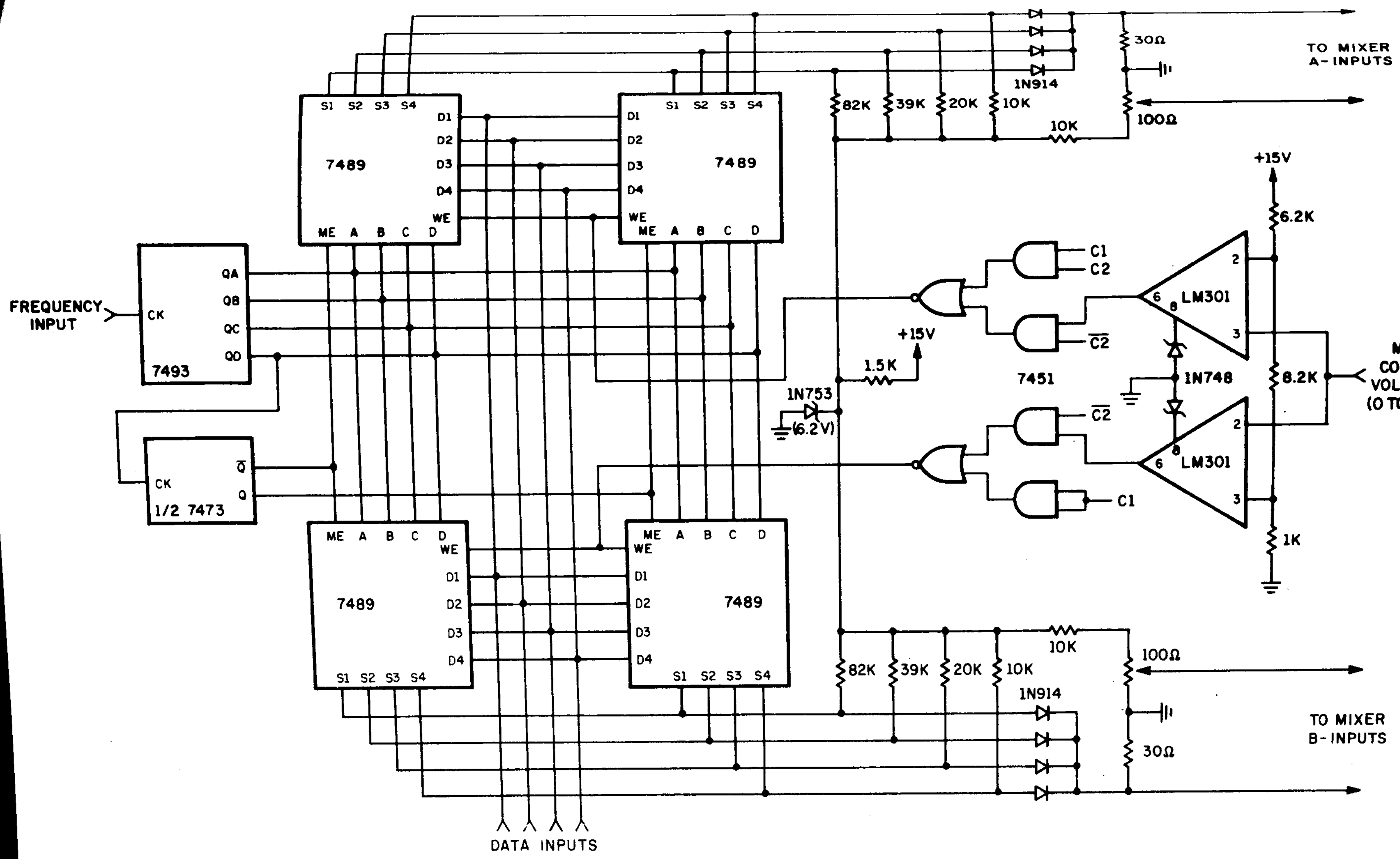


Figure 10. Digital Waveshape Generator.

introduced by the D/A converters and also to compensate for possible offsets caused by mismatches between the input transistors of the mixer.

Figure 10 also shows the circuitry involved in the control of memory read/write operations. The pair of LM 301 comparators with TTL-compatible outputs constitutes a window detector having thresholds of 1V and 9V respectively. The detector continuously monitors the mixer programming voltage w and allows either pair of memories to be updated only when the corresponding output is under complete attenuation by the mixer. Input v_A is attenuated when w is above 9V and input v_B when w is below 1V. When w lies inside the window, neither signal is completely attenuated and memory updating is inhibited. This mode of operation is particularly desirable for the production of continuous, smooth waveshape changes and is selected by setting control bits C1 and C2 both to logic 0.

When C1C2 = 01, memory-write operations are inhibited altogether, regardless of the window detector response. This mode is useful for amplitude modulation, as will be discussed in more detail in connection with the mixer. Such an effect is implemented by programming the mixer with a modulating signal which causes the output waveshape to glide repetitively back and forth between the two waveforms trapped in the respective memories.

When C1C2 = 10, the read/write mode of the upper memory pair is controlled by the corresponding comparator in the usual way, while the other comparator is overridden and the lower memory pair is permanently forced into the write mode. This control state offers the option of bypassing one of the memory pairs, so that the information present at the data inputs can be injected directly into the mixer, via the D/A converter, where it is then combined with the waveshape emerging from the other memory pair. This option is made possible by the fact that the SN7489 memories, when the write mode

is enabled, display at the sense outputs the same information--albeit in complemented form--that appears at the data inputs.

The usefulness of this control state is typified by the generation of pitched noise, which is easily achieved by driving the memory data inputs with a pseudo-random, binary number generator and by setting C1C2 = 10. The presence of the mixer allows for the programming of arbitrary amounts of pitching, while the output low-pass filter can be exploited to impart different types of coloration to the sound emerging from the mixer. The class of sounds thus obtained is particularly useful in the synthesis of percussive sounds of the crash-cymbals or snare-drums type, or in the creation of special effects like steam gushings, surf surges, etc.

The control state C1C2 = 11 puts both memory pairs into the write mode and causes them to be loaded with the same information. This mode is useful for amplitude modulation with carrier suppression, as will be discussed in more detail in connection with the mixer.

The information to be loaded into the memories during write-operations is normally derived from a central, pseudo-random, binary number generator. Since the amount of hardware and control is kept to a bare minimum, this form of information generation offers the significant advantages of simplicity and economy. It also offers a more unified approach to sound synthesis in that the same circuitry can be used for the generation of harmonic as well as noise-like sounds. Although this method, in conjunction with the control options illustrated above, has proven quite adequate to span a vast range of improvisational situations, there are circumstances which ask for a strictly deterministic approach to the generation of waveshape information. To handle these situations, a central memory bank is provided, where the information pertaining to the waveshapes of interest is stored and made available for

consumption by any of the harmonic generators. Thus, additional steering and control facilities are provided for bypassing the binary noise-generator and for properly routing the desired information.

A typical example of this class of situations is offered by the synthesis of speech-like sounds. Being harmonic to a considerable degree, human vowels are particularly suited to the digitized synthesis technique here described. Since each vowel has a predetermined waveshape, the associated digital information is stored in a central read-only memory from where it can be loaded, upon request, into any of the peripheral waveshape generator memories. It is worth pointing out, in passing, that the presence of the programmable mixer offers a rather simple means for the generation of diphthongs, and that the control options associated with pitched and colored noise can be put to good use for the synthesis of a number of fricative consonants.

2.6 Programmable Mixer/Modulator.

The function of this circuit is to mix two input signals v_A and v_B in complementary percentages, with the mixing ratio being programmable by an external control voltage w. Thus, if v_0 denotes the mixer output, the desired transfer function is

$$v_0 = (w/10\ V)\ v_A + (1 - w/10\ V)\ v_B,$$

where control-voltage w spans the range 0 to +10 V.

As shown in Figure 11, the mixer is centered around two variable-transconductance transistor pairs, very much like present day four-quadrant IC multipliers.[13,14] However, unlike in a multiplier, the inputs to the transistor pairs are kept separate and their collectors are connected in phase. The bias current for powering the transconductance elements is provided by current generator Q8. After going through differential pair Q3/Q6, this current is split into two complementary components which bias

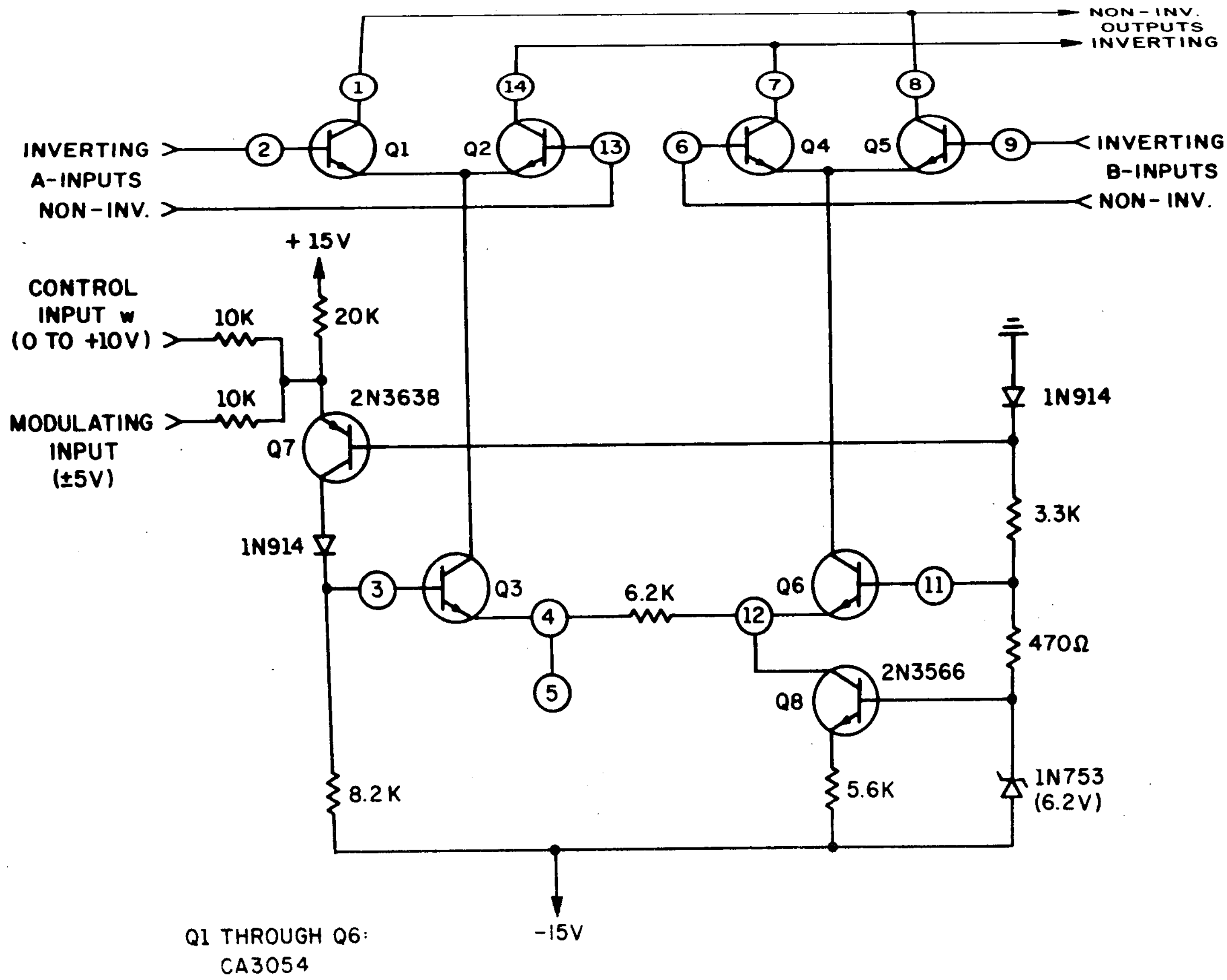


Figure 11. Programmable Mixer/Modulator.

transconductance pairs Q1/Q2 and Q4/Q5 independently. As the gain of a transconductance element is proportional to its bias current, the relative contributions of the two pairs to the common output are determined by the ratio of the respective bias currents. Hence, the mixing ratio can be easily controlled by programming current splitter Q3/Q6.

The circuit parameters have been chosen so that a control-voltage span of 10 volts slightly overdrives the current splitter. This precaution ensures that at the edges of the control range one of the transistors of the current splitter is completely cut off and the corresponding transconductance element brings no contribution to the mixer output. The transitions from linear to saturating regions occur around 1 V and 9 V respectively, and take place smoothly by virtue of the knee exhibited by the base-emitter junction characteristics of Q3 and Q6.

The nature of the input stage of the programmable low-pass filter, to be discussed below, allows for the balanced output lines of the mixer to be fed directly to the filter with no need for additional interface circuitry.

Owing to the inherent multiplicative capabilities of the mixer, the areas of application of this device can be further expanded to include various forms of amplitude modulation. To this end, in addition to the input for control voltage w, another terminal is provided for applying the modulating signal. Both inputs are then added together by common-base transistor Q7.

When the circuit is operated as a modulator, waveshape updating is usually inhibited (C1C2 = 01), owing to the fact that the purpose of the window detector is defeated by the presence of the modulating signal. The type of modulation most commonly obtained is without carrier suppression, since the circuit is basically a two-quadrant multiplier. However, the carrier can be suppressed by having $v_A = -v_B$ and by d.c. offsetting the modulating signal by 5 V, as can be readily verified from the mixer transfer function.

These conditions are achieved by first setting control state C1C2 = 11, which causes both memory pairs to be loaded with the same waveshape. Subsequently, control bit C1 is changed to zero, while w is being set to +5 V. The required phase inversion of one of the waveshapes is ensured by the fact that the balanced outputs from the corresponding D/A converter are wired interchanged to the mixer. This asymmetry does not disturb the other control modes since the phases with which waveshapes are loaded into memories are immaterial.

III. ELECTRONICALLY PROGRAMMABLE FILTERS

A filter is electronically programmable if its characteristics--corner or center frequency, damping factor, etc.--can be altered by means of external control signals. Although a considerable amount of work in the area of active filters of fixed or manually programmable characteristics has been going on for more than a decade now,[15] electronically tunable filters have made their appearance only in very recent years. Filter parameter control is achieved by means of devices exhibiting some form of multiplicative capabilities. JFET voltage-controlled resistors,[16] analog multipliers,[17,18] and periodic-switched filter networks with variable duty-cycle[19,20] are examples of the multiplicative components and techniques being presently used. All of these methods, however, seem to suffer from one kind of limitation or another which restrict their useful ranges of operation to no more than a few octaves.

The need for a filter of much wider programming range has led to the development of two building blocks which, when properly utilized in the synthesis of filters, allow for more than three decades of parameter control. A range of this width is made possible by the fact that the controlling signals are currents, rather than voltages. Indeed, currents can be used to accurately represent analog variables over much wider ranges than voltages, a feature of present-day technology which never seems to be stressed enough. In fact, by exploiting the highly predictable relationship between collector current and base-to-emitter voltage drop of the silicon transistor, currents can be easily controlled over a number of decades,[11] while ordinary voltage ranges are dynamically limited by temperature drift and other forms of noise.

The heart of the filter blocks in question is a programmable transconductance element, also referred to as an operational transconductance

amplifier, or OTA for short. The device resembles an ordinary operational amplifier in that is also has one output and two input ports, the polarity of the output signal being the same as that of the voltage at one of the input terminals, and opposite to that of the other. However, unlike operational amplifiers, the OTA's output signal is a current, rather than a voltage, and the transfer characteristic--which has therefore the dimensions of a conductance--can be controlled externally by means of a current signal, also referred to as a bias current I_{BIAS}. Thus, in addition to the three signal ports mentioned above, the OTA has a fourth control port for the application of I_{BIAS}. The definition of the ideal OTA is illustrated in Figure 12.

The filter building block to be described first is provided by the circuit of Figure 13. Using Laplace transforms,

$$V_o = I_o/(sC) = g_m(V^+ - V^-)/(sC)$$

or

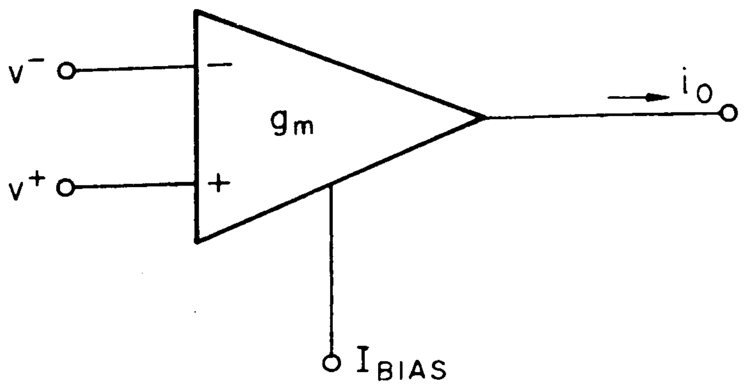
$$V_o = (V^+ - V^-)/s\tau$$

where

$$\tau = C/g_m = (C/K)/I_{BIAS} \cdot$$

Thus the transfer function of the circuit is that of a bipolar integrator, with the constant of integration being externally programmable by means of I_{BIAS}. This simple building block can be readily incorporated into filter configurations based on integrators, like the state-variable[21] or the biquad[22] topologies. Then, by controlling I_{BIAS} with an exponential current generator, the constant of integration, i.e. the natural frequency of the filter, can be readily varied over a wide dynamic range.

As an example, Figure 14 shows how two such integrators can be connected together to synthesize a filter which exhibits simultaneously a low-pass and a band-pass transfer function of constant bandwidth. By working with Laplace transforms in the manner indicated above, one can readily verify that



$$Z_i^+ = Z_i^- = Z_o = \infty$$

$$i_o \triangleq g_m \, (v^+ - v^-)$$

$$g_m \triangleq KI_{BIAS} \;, \quad K = \text{constant}$$

$$[g_m] = \Omega^{-1} \;, \quad [K] = V^{-1} \;.$$

Figure 12. Definition of Operational Transconductance Amplifier (OTA).

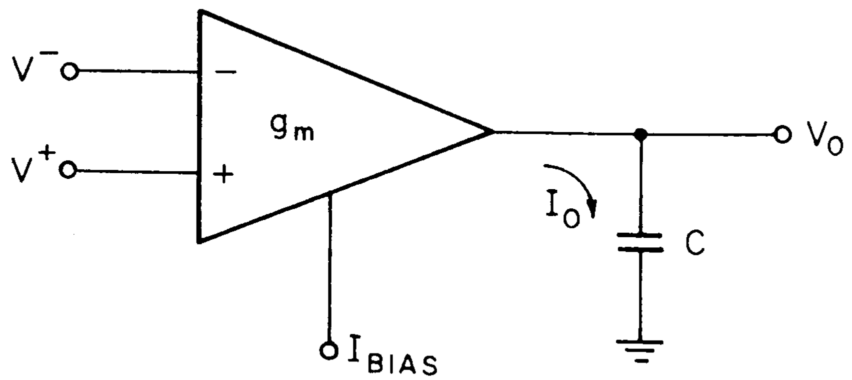


Figure 13. Programmable Bipolar Integrator.

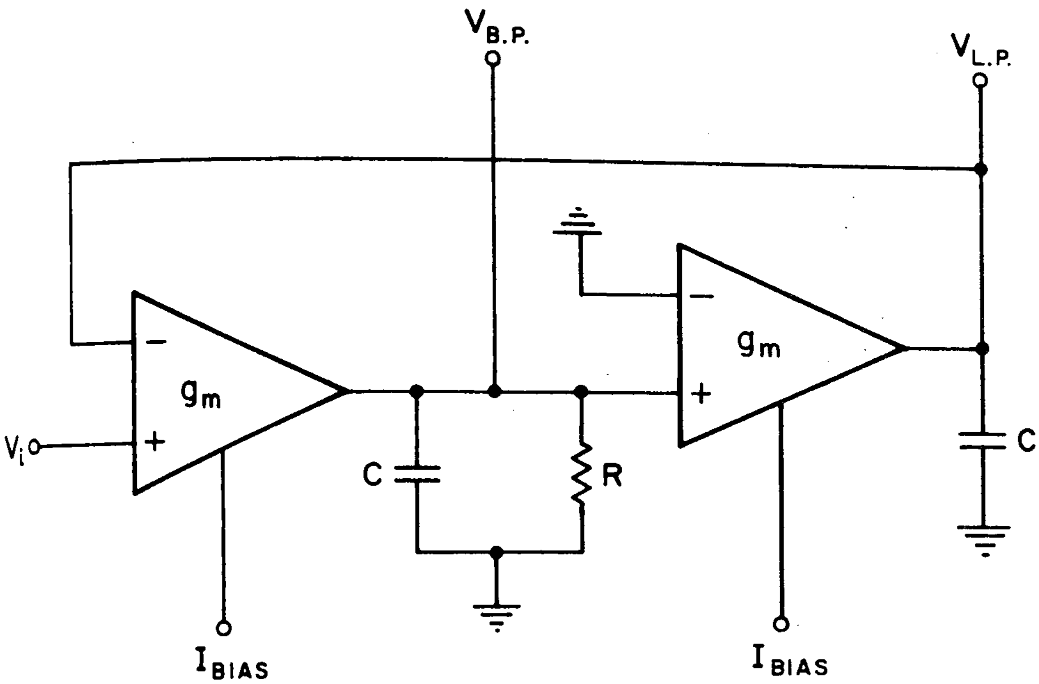


Figure 14. Two-integrator Programmable Filter.

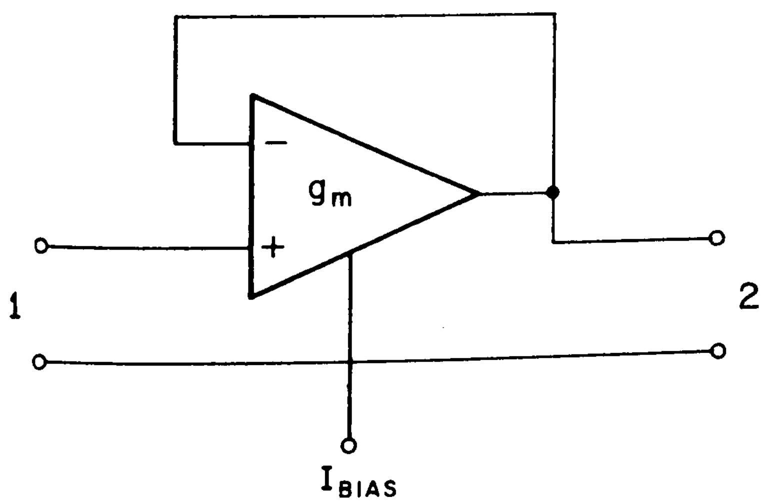


Figure 15. Programmable Conductance.

$$V_{B.P.}/V_i = s\omega_o/(s^2 + s\omega_o/Q + \omega_o^2)$$

and

$$V_{L.P.}/V_i = 1/(s^2 + s\omega_o/Q + \omega_o^2) ,$$

where

$$\omega_o = 1/\tau = KI_{BIAS}/C$$

is the natural frequency of the filter, and

$$\omega_o/Q = 1/(CR)$$

is the constant bandwidth of the band-pass function. Resistor R has been added to insert a loss and therefore ensure a non-zero damping factor. If R is omitted altogether, the circuit oscillates and, as such, it can be used as a programmable sine-wave oscillator with quadrature outputs. Conversely, the replacement of R with a programmable conductance, to be discussed below, offers the option of independent control over the Q or over the bandwidth of the filter.

The other building block is offered by the feedback configuration of Figure 15. The behavior of this circuit can be concisely described by means of its y-parameters, which are

$$y_{11} = y_{12} = 0, \quad y_{21} = -g_m, \; y_{22} = +g_m .$$

However, its usefulness as a building block for the synthesis of programmable filters can be better visualized by considering its equivalent circuit, which is shown in Figure 16. The circuit consists of a current-controlled conductance in series with a voltage follower to account for the infinite impedance as seen into port 1 of the original circuit. If two devices like the one of Figure 15 are connected in parallel but with the ports interchanged, the resulting circuit behaves like a programmable, floating conductance. Hence, it is a simple matter to replace some or all of the resistors of a given filter with programmable conductances and thus make the filter itself programmable.

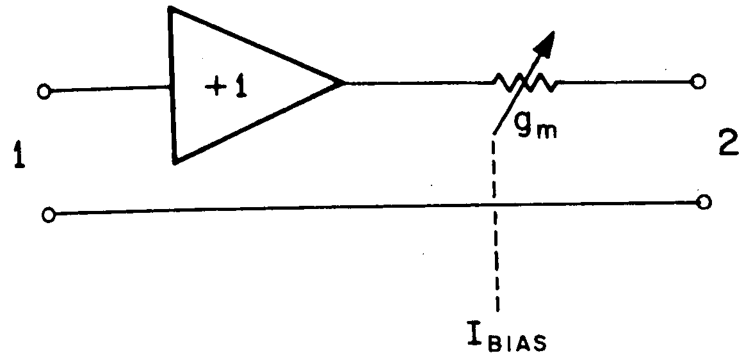


Figure 16. Equivalent Circuit for Figure 10.

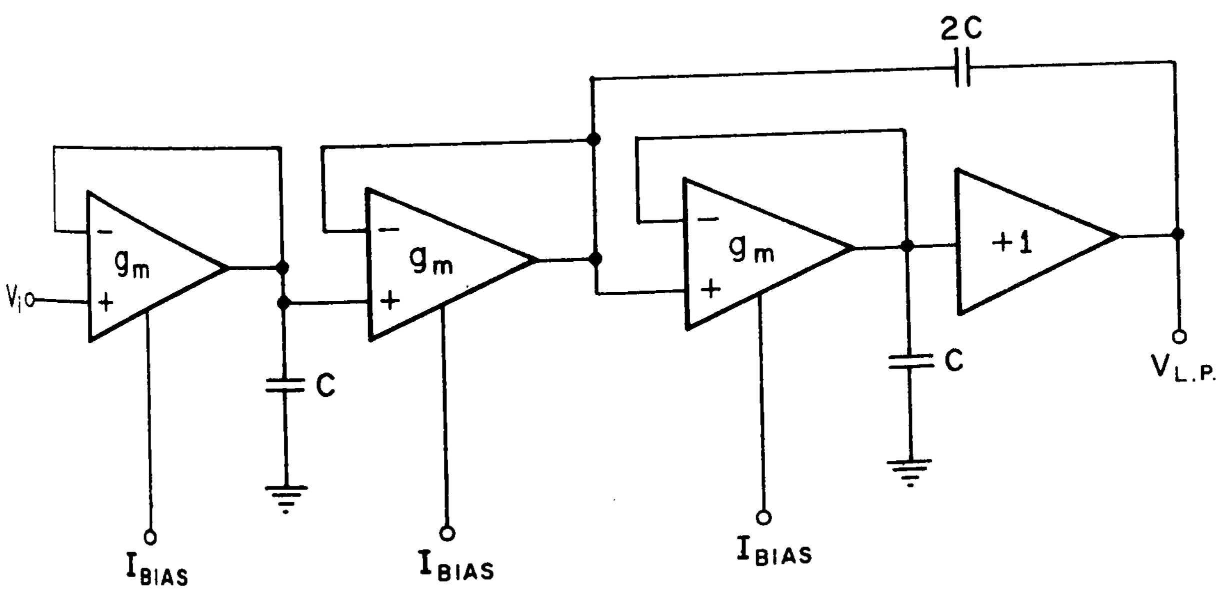


Figure 17. Three-pole, maximally-flat, low-pass programmable filter.

In most cases of interest, however, a single OTA per resistor is sufficient, without the need for the parallel configuration just discussed, which requires two OTA's instead of one. Indeed, the high input-impedance presented by port 1, far from being a limitation or a nuisance, is often a desirable feature since it facilitates network coupling, and it can also be exploited to improve filter sensitivity.[23]

Figure 17 shows one way of realizing a three-pole, maximally-flat, low-pass programmable filter with three conductance building blocks of the type just described. The filter uses also a unity-gain voltage follower to decouple the output of the third OTA from the feedback capacitor and outside loads.

Figure 18 shows a circuit realization of the operational transconductance concept which could readily be put in integrated form since it involves circuit elements of well established feasibility in commercial linear IC's. The input stage, as in present-day analog multipliers,[14] consists of a balanced voltage-to-current converter (Q1 through Q4) and a diode compressor (Q5/Q6) to provide the appropriate base-drive for transconductance pair Q7/Q8. This configuration results in a substantial improvement for the linearity of the transconductance pair, while drift and other noise factors are also minimized.[24] Because of its Darlington-connected, differential amplifier configuration, the voltage-to-current converter presents a high input-impedance which allows for an OTA to drive other OTA's directly, as shown in Figures 14 and 17.

The balanced collector currents of transconductance pair Q7/Q8 are converted into a single-ended difference output current by means of current-mirrors Q9/Q10/Q11, Q12/Q13/Q14, and Q15/Q16/Q17. To achieve high output-impedance, these current-mirrors make use of the Wilson configuration.[25]

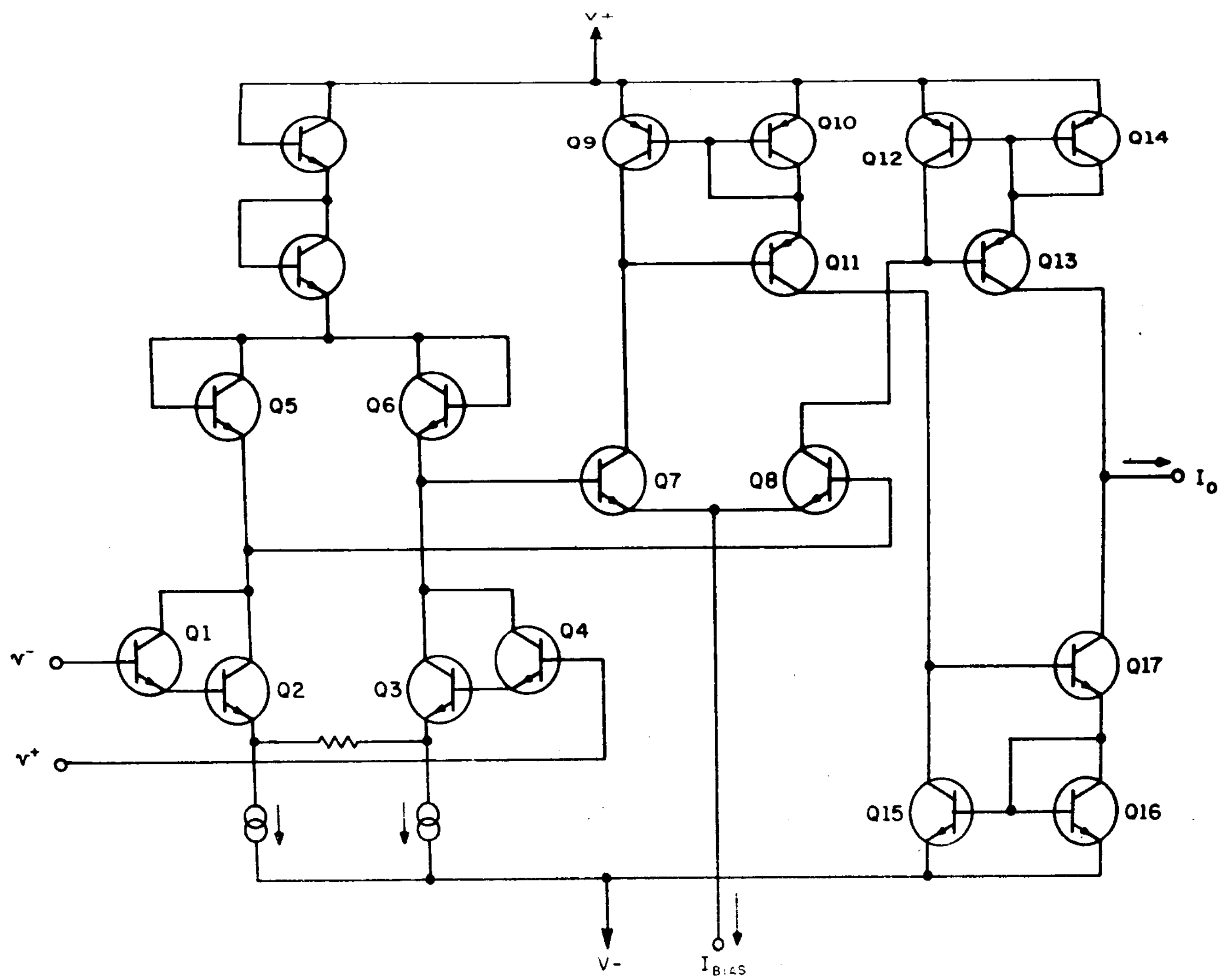


Figure 18. Operational Transconductance Amplifier.

The value of K is established by the circuit parameters of the voltage-to-current converter--bias currents and amount of emitter degeneration. A convenient parameter choice which conforms with present analog standards is the one which results in $K = 1/(10\ V)$.

Although the complete circuit of Figure 18 is not available in integrated form, an IC OTA of similar concept but without the input converter and compressor is presently manufactured by RCA (CA3080). For this OTA implementation,[26] $K = q/2kT$, so that, at room temperature, $K = 19.2 \cdot 10^{-3} V^{-1}$. The CA3080 also requires that I_{BIAS} flow <u>into</u> the current control terminal which is kept at a negative potential.

The smoother for the harmonic tone generator has been realized with two CA3080 OTA's in the manner illustrated in Figure 19. This circuit realizes a two-pole, maximally-flat, low-pass programmable filter. Due to the absence of the voltage-to-current converter and diode compressor, signal conditioning at the OTA's inputs is achieved by means of resistor attenuators, as shown. The resistor values have been chosen so that the balanced outputs from the programmable mixer can be tied to the inputs of the first OTA directly, without the need for additional ancillary circuitry. Also, because of the finite impedance presented by the attenuators, the output from each OTA must be buffered by a high input-impedance decoupler. This task is performed by the pair of FET-transistor source followers. The voltage-offsets associated with the followers do not affect the dc level of the filter because each follower is placed inside the feedback loop of the corresponding OTA.

Also shown in Figure 19 is a dual exponential current generator which provides the appropriate bias for both OTA's. The generator consists of three matched pnp transistors from the CA3084 array and a discrete, high-beta npn transistor. Being forced to conduct a constant current--about 1 mA--

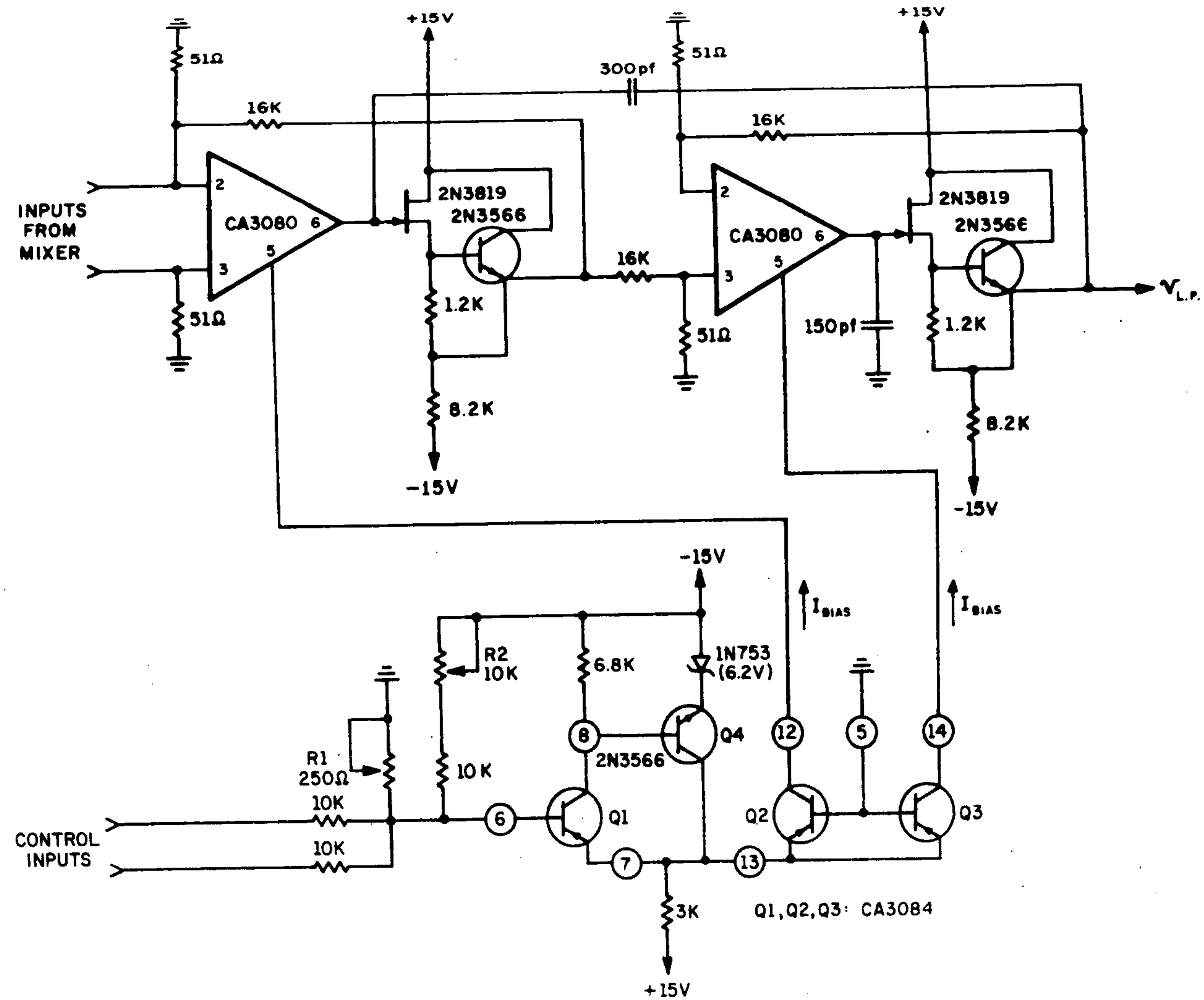


Figure 19. Two-pole, Maximally-flat, Low-pass Programmable Filter.

by regulator Q4, transistor Q1 develops a base-to-emitter voltage drop which, apart from temperature fluctuations, is also constant. Thus the emitter terminal of Q1 constitutes a low-impedance point and presents a fixed voltage-offset with respect to the base terminal. This is just what is needed to provide the proper emitter drive for exponential current generators Q2 and Q3. Since the emitters of Q1, Q2, and Q3 are tied together, any voltage excursion applied to the base of Q1 is transmitted to the emitters of Q2 and Q3 unaltered, by virtue of the regulating action of Q4. Furthermore, temperature variations of the base-to-emitter voltage drop of Q1 tend to cancel out analogous variations affecting Q2 and Q3, thereby providing temperature stabilization for the exponential generators.

Trimpot R2 sets the base and trimpot R1 the width of the exponential range. One of the exponential generator inputs is fed with the Pitch Voltage control signal, so that the position of the filter corner-frequency, relative to the sampling frequency, is kept constant over the whole range of interest. The other control input is used to perturb or alter this fixed constraint, thereby providing an additional form of spectral control besides the techniques previously discussed.

IV. SOUND INTENSITY AND LOCATION

4.1 Intensity and Location Control.

The circuitry described so far has dealt primarily with the synthesis and control of pitches and timbres. Another parameter of considerable psychoacoustical consequence is intensity. As in the case of tonal color, particularly important in music are the dynamic characteristics of this parameter, also referred to as the sound envelope. These characteristics are related to the manner in which sounds build up and decay in time, and constitute a significant clue toward the instrumental characterization of sounds. Given the importance of this parameter, it is desirable that intensity control allow for a broad choice of sound envelopes in order to ensure a wide spectrum of instrumental capabilities.

In analog systems, the dynamic control of sound intensity is realized by means of programmable attenuators operating under the control of envelope generators. The types of envelope functions usually available are triangular, trapezoidal, or trapezoidal-with-overshoot, and such envelope parameters as attacks, steady-states, and decays must be individually preset by hand. These techniques, far from satisfying a broad range of instrumental situations, compound real time performance in that they necessitate manual control. By using hybrid techniques, as it will be shown below, both of these problems can be approached in a more organic fashion that allows the control of complex and minute deviations in a manner that parallels the domain of natural sounds.

Another parameter that has become rather significant in the music of the last two decades is the location and movement of sound within a musical environment. For studio-type operations, the output of the system under discussion

consists of four discrete channels. The control of spatial parameters is accomplished by dynamically programming the distribution of sounds among the four channels, and by adding programmable amounts of artificial reverberation to the original sounds.

In hardware terms, sound enveloping and quadraphonic sound location are handled uniformly by means of a programmable attenuator/locator circuit. This device acts simultaneously as programmable attenuator and programmable joy-stick.

4.2 Programmable Attenuator/Locator.

Since the harmonic tone generator produces sounds of constant amplitude, signal level control is realized independently by feeding the generator output into a programmable attenuator/locator and by controlling the amount of attenuation with an envelope generator. A block diagram of the attenuator/locator is shown in Figure 20.

Owing to the fact that the human ear responds to the amplitude of sounds in a nonlinear manner, it is necessary that the linear domain of the envelope voltage first undergo an appropriate nonlinear mapping. A mapping which is musically adequate as well as electronically convenient is the exponential one, and the related circuitry is shown in Figure 21. The exponential conversion proper is performed by transistor Q1 while transistor Q3 and regulator Q2 serve the purpose of providing the appropriate emitter drive for Q1. Thus the circuit is of the type already encountered, except for a certain degree of simplification justified by the less stringent requirements of level control. Circuit parameters have been chosen so that a 0 to 10 V control input range is mapped into an exponential current range of about 50 db. Although the circuit can easily afford a much wider range, 50 db constitutes a convenient compromise that ensures adequate musical range without imposing

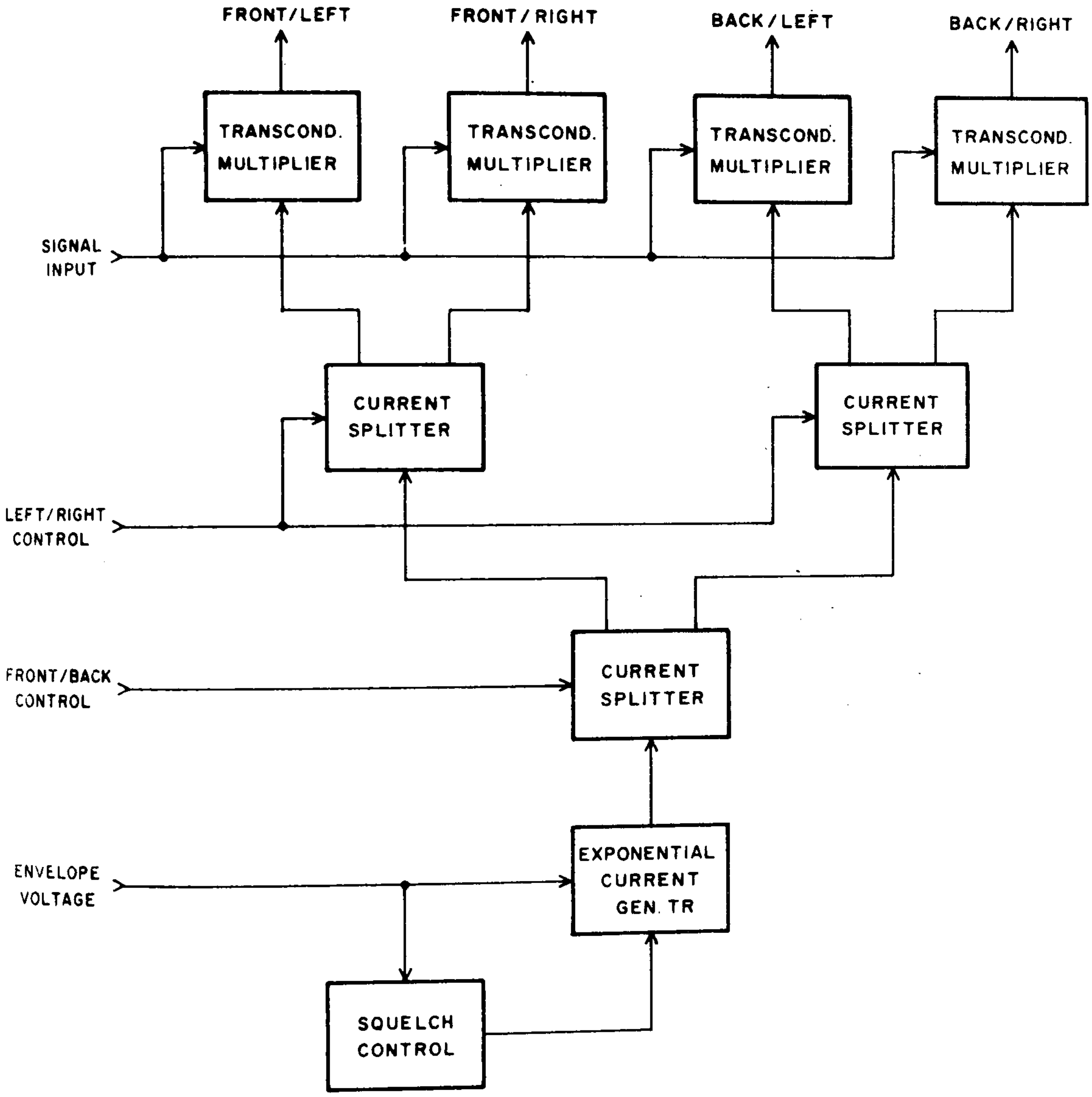


Figure 20. Block Diagram for Programmable Attenuator/Locator.

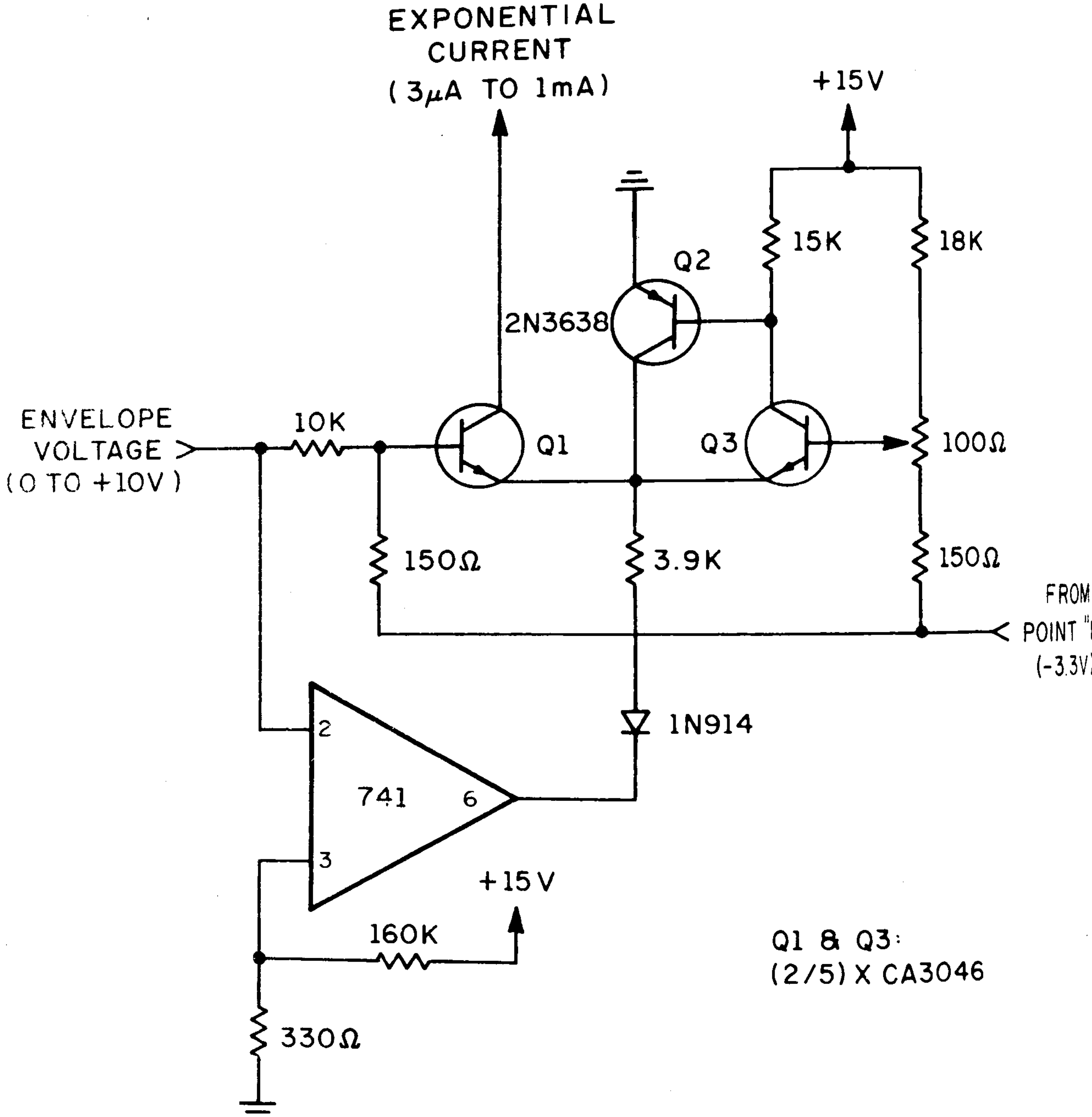


Figure 21. Exponential Current Generator and Squelch Control.

too stringent demands on envelope control. However, 50 db below full signal level does not offer enough attenuation when silence is desired. This inconvenience is eliminated by means of a squelching circuit which senses the input control voltage and completely shuts off the exponential generator as soon as the input goes below a few tens of millivolts. This artifice greatly expands the total effective width of the range, while still keeping the useful control portion at about 50 db.

The exponential current thus obtained is subsequently split into four separate components which are used to bias four corresponding variable-transconductance transistor pairs, one for each output channel. The signal input terminals to all four pairs are tied together. Since the gain of a transconductance pair depends on the amount of emitter bias current, it is immediately seen that the output signal level from one of the pairs, relative to the levels of the others, depends on the percentage of exponential current diverted to that pair. Furthermore, all four gains depend on a common scale factor proportional to the magnitude of the exponential current. Thus the exponential generator determines the overall sound intensity and the current splitters determine the relative intensity distribution among the four channels.

The circuit schematic for the transconductance pairs and current splitters is shown in Figure 22. As in the case of the programmable mixer, the choice of circuit parameters is such that the current splitters are slightly overdriven by a control voltage of 10 V range. This precaution ensures that at each end of the control range one of the current splitting transistors is totally shut off, thereby providing full signal attenuation at the transconductance pairs biased by that transistor. This feature allows complete channel isolation.

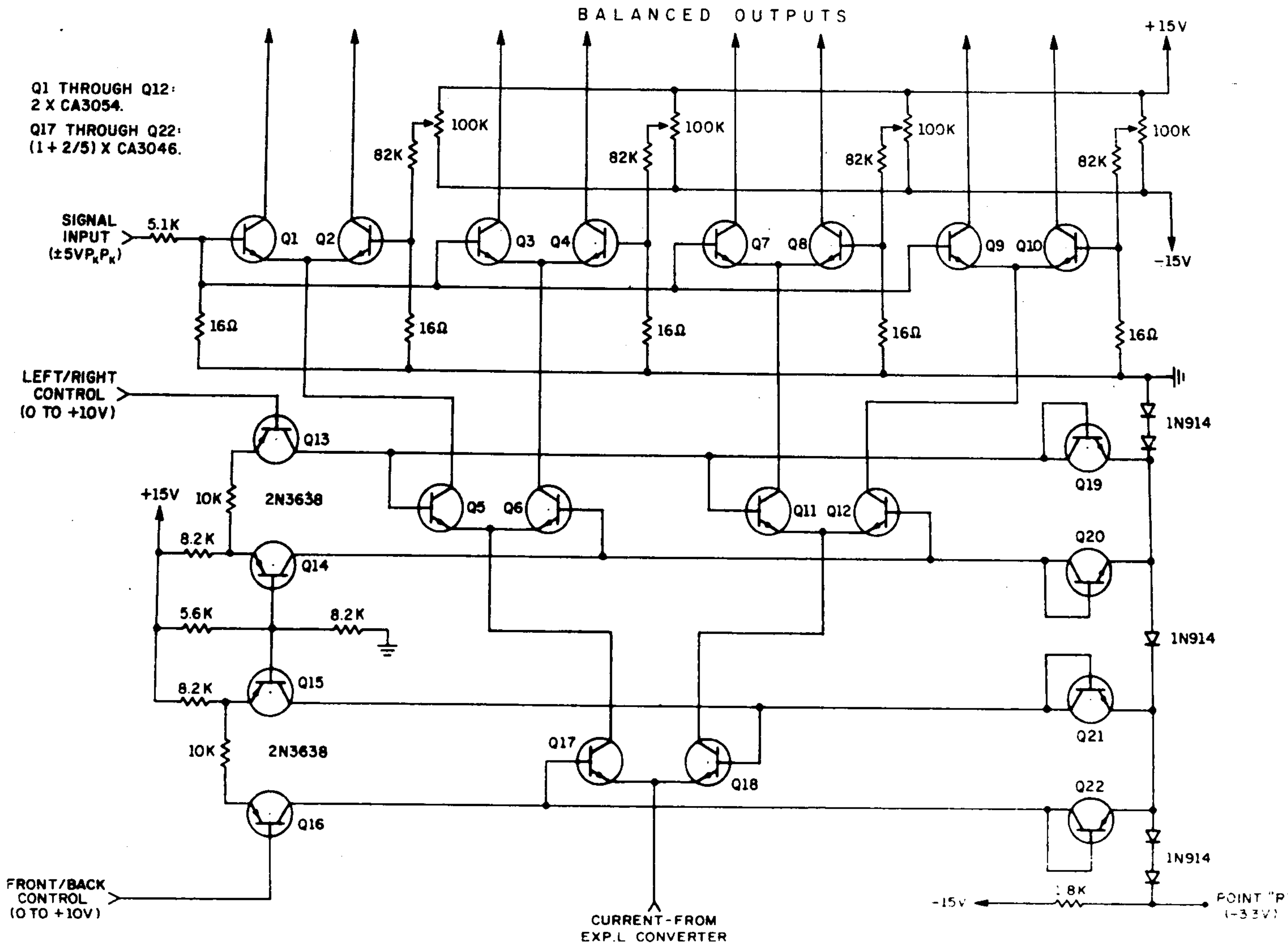


Figure 22. Programmable Attenuator/Locator.

Diode-connected transistor pairs Q19/Q20 and Q21/Q22 constitute a non-linear voltage compressor whose purpose is the linearization of the current splitters' transfer characteristics, a topic already discussed in connection with OTA's. The proper bias voltages for the current splitters and the exponential current generator are provided by the string of diodes shown at the right of the circuit diagram.

To illustrate the above arguments with an actual example, consider transistor pair Q17/Q18, which constitutes the front/back current splitter. When the front/back control voltage is around OV, Q15 is cut off, so that Q21 and Q18 are also off. Thus all the exponential current is diverted to the "front" transconductance pairs, while the "back" pairs are completely cut off. If the control voltage is now increased, conduction is gradually transferred from Q16 to Q15. This is accompanied by a simultaneous increase in the conduction of Q21 and Q18 and decrease in the conduction of Q22 and Q17. Thus, the exponential current is gradually diverted from Q17 to Q18, causing a gradual attenuation at the "left" channels and gradual amplification at the "right" channels. Upon reaching 10 V of control voltage, Q16, Q22 and Q17 are completely cut off, and the situation is now reversed, as compared with the case in which the control voltage is around OV. Similar arguments hold for the left/right current splitters.

It is a known fact that transconductance pairs, particularly when implemented with general purpose IC transistors, generate an undesirable amount of noise which limits their usage in high-quality audio applications. Among others, the noise level depends on the amount of emitter bias, and usually increases with the latter. This effect, however, is not so critical in the present application, owing to the fact that the attenuator/locator is being permanently fed with a full-level signal from the tone generator.

Thus, an increase in emitter bias increases not only the noise level, but the signal level as well, thereby ensuring a satisfactory signal-to-noise ratio over the entire dynamic range of interest. Subjectively, this ratio is further improved by the well known masking effect of human hearing, whereupon weak sounds--noise in this case--are partially masked by loud sounds in the same frequency band.

The circuit realization of the entire attenuator/locator, inclusive of the exponential generator, requires four IC transistor arrays, besides the ancillary circuit shown. Two CA3054 IC's are needed to implement the four transconductance multipliers and the left/right current splitters. One CA3046 implements the remaining splitter and the exponential generator, and one more CA3046 is needed for the diode compressors. Each transconductance transistor pair is provided with a trimmer for input offset compensation.

To render the attenuator/locator suitable for driving standard audio equipment, the balanced outputs from each transconductance pair must be converted into a single-ended voltage signal. This is readily accomplished by means of a differential operational amplifier, in the conventional manner of analog multipliers. However, since all eight voice modules share the same four output channels, it is not necessary to use an operational amplifier for each of the 32 transconductance pairs. Rather, the outputs from all eight voice modules are connected in parallel, with corresponding collectors of corresponding transconductance pairs being tied together. The desired double to single-ended conversion is then accomplished with four differential amplifiers only. This circuit configuration results in the saving of 28 operational amplifiers.

Before being fed to output audio equipment, each channel is sent through a spring reverberator and the spring output is mixed in programmable amounts

with the unreverberated signal. The associated hardware, involving standard audio circuitry, will not be described here any further.

4.3 Envelope Generation.

To control intensity dynamically, the exponential generator discussed above must be programmed with control functions of time, or envelopes. This task is handled in hybrid fashion by means of an envelope generator in conjunction with a binary word sequence generator.

Envelopes are generated in the form of joined straight line segments. Thus the characteristics of a given envelope are entirely defined once the locations of its breakpoints are known. The specification of a breakpoint involves eight bits of information. Four bits of amplitude information determine the breakpoint ordinate, and four bits of rate information determine the slope of the line segment terminating at that point.

The typical generation of a segment occurs as follows. The envelope generator sends the binary word sequence generator a request for an 8-bit word of data. After receiving the data, the envelope circuit generates a voltage ramp at the rate specified by the homonymous bits. The polarity of the ramp is implied by the remaining four bits of amplitude. Depending on whether these bits specify a greater or a smaller voltage than that currently output by the envelope generator, the ramp polarity is respectively positive or negative. Thus, in the course of ramp generation, the output voltage approaches linearly the value specified by the amplitude bits. Upon reaching this value, the generation of the current line segment is completed, and a new data request is issued for the generation of another segment.

Four bits of amplitude as well as rate resolution may seem inadequate to handle detailed envelope characteristics. It should be noted, however,

that 16 loudness levels over a dynamic range of 50 db yield a resolution of about 3 db, which compares favorably enough with the sound loudness resolution of the human ear. If a fine control over the envelope structure is desired, this can be simply achieved by expressing the envelope function in terms of small line segments.

The envelope generation process is illustrated in the block diagram of Figure 23. The heart of the circuit is an operational amplifier with programmable slew-rate. This device is readily simulated with the help of two ordinary amplifiers: a comparator to simulate the input stage, and a current integrator to simulate the Miller stage. Slew-rate control is achieved by programming the current feeding the integrator. To ensure a wide dynamic range of envelope characteristics, the current is generated by means of an exponential converter, as in the control of pitch and loudness.

When the output from the amplitude D/A converter exceeds the output from the integrator, the comparator sets the current switch to the "0" position. The integrator is therefore driven from the exponential generator directly and produces a positive ramp, or attack. When the integrator output exceeds the converter output, the comparator sets the switch to the "1" position. Being now diverted to the current mirror, the exponential current undergoes a polarity reversal, thereby causing the integrator to produce a negative ramp, or decay. Thus, the chain comparator-switch-integrator constitutes a voltage follower with externally programmable slew-rate. The closed-loop frequency parameters of the chain are such that, when the output becomes equal to the input, the comparator begins to oscillate between the "0" and "1" states. This feature is intentionally exploited to mark the achievement of an envelope breakpoint and therefore issue a request for a new word of data from the binary sequence generator. The request generation is handled by the breakpoint detector circuit.

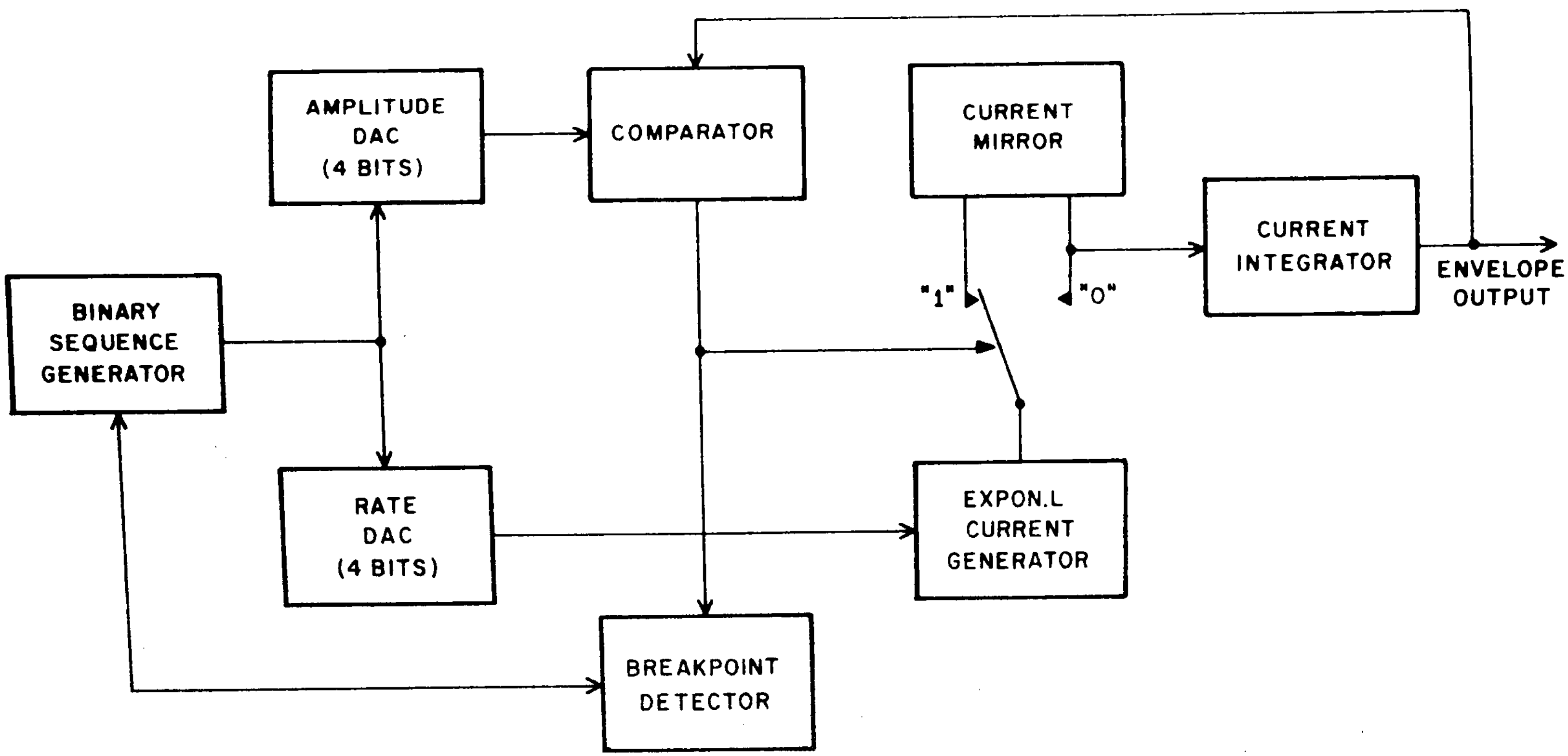


Figure 23. Envelope Generation Block Diagram.

The detailed circuit diagram of the envelope generator, exclusive of the binary sequence generator, is shown in Figure 24. Binary data are separately converted into analog signals by means of diode-resistor networks and transistor switches, as shown. The rate D/A converter is fed to the exponential generator, which consists of transistors Q4, Q5 and Q6 connected in the usual manner. The amplitude D/A converter is fed to the TTL-compatible comparator where its output is compared against the output from the integrator. Depending on whether the amplitude of the converter output is smaller or greater than that of the integrator, the exponential current is diverted to the integrator input either through diode D1 or through current mirror Q1/Q2. Transistors Q1 through Q5 belong to the CA3096 IC array. Upon reaching an envelope breakpoint, the comparator begins to oscillate, causing the 2-bit counter to overflow and thereby generate a request pulse to the binary sequence generator. When servicing the request, the sequence generator also clears the counter, thereby providing proper initialization for the detection of the next breakpoint.

With the circuit parameters shown, attack slew-rates cover a range from 1 V/ms to 1 V/s and decay slew-rates from 2 V/ms to 2 V/s. The one-octave discrepancy between the two ranges is dictated by reasons of musical convenience and is handled by offsetting the exponential range by a corresponding amount in the course of decay ramps. The offsetting circuitry is centered around resistor R1 and transistor switch Q7. During an attack the comparator output is at "0" and the offsetting circuitry is disabled by switch Q7. During a decay the comparator opens switch Q7, so that the current through resistor R1 is directed to the base of exponential generator Q4, where it introduces a voltage offset of about -18 mV. This causes the collector current of Q4 to decrease by one octave, as desired. Trimmer R2 serves the purpose of adjusting the base of the exponential range.

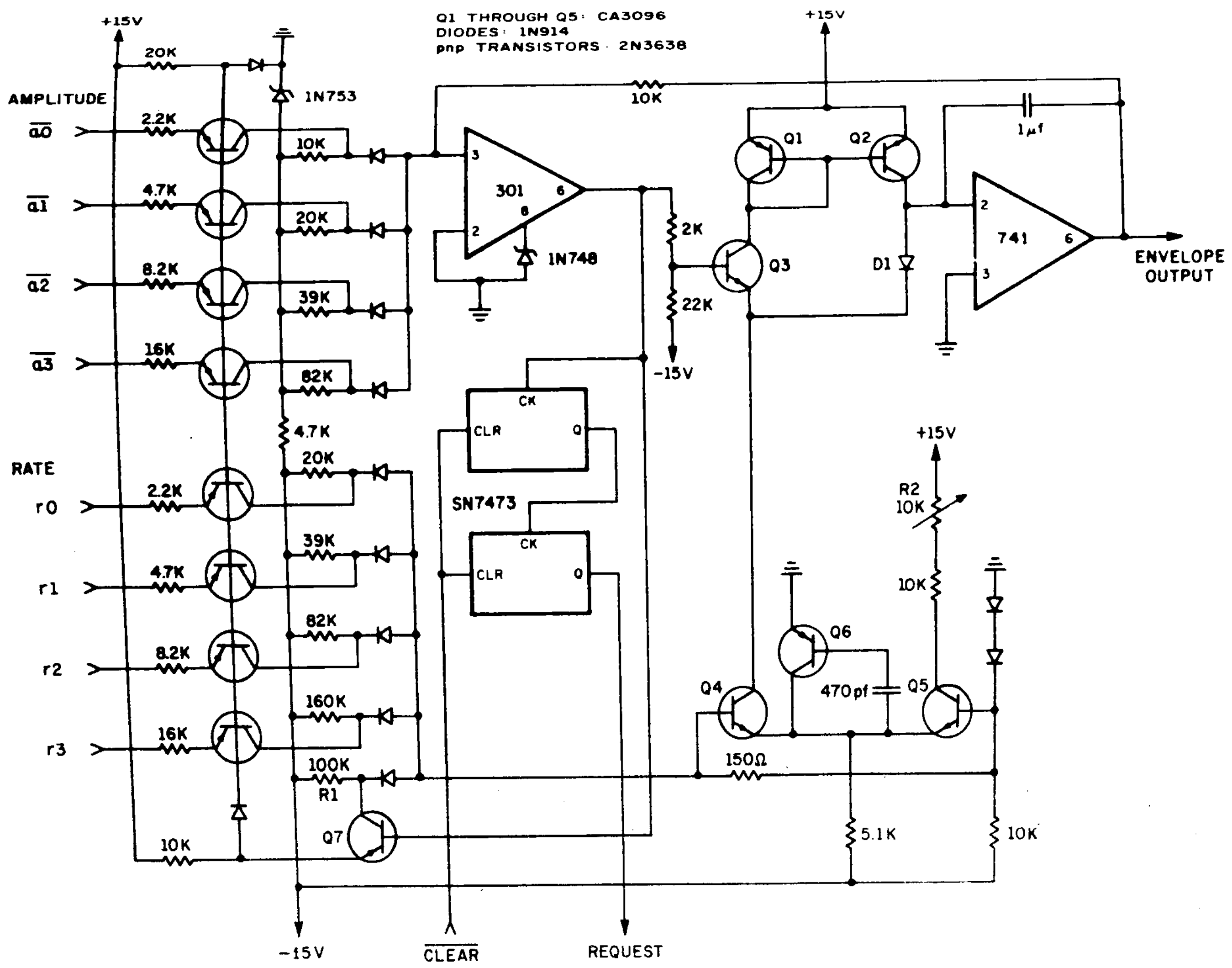


Figure 24. Envelope Generator.

As will be seen in connection with musical parameter control, the applications of piece-wise linear function generators are not confined to sound enveloping alone. When used in applications that do not require the one-octave discrepancy between positive and negative slope ranges, the offsetting circuitry just described is simply omitted from the hardware realization of the generator.

V. PERCUSSION ENSEMBLES

Two of the four orchestras, in addition to the voice modules described above, include also a 16-element percussion ensemble each. In principle, percussion-like sounds can be synthesized by means of voice modules, as all other sounds are. However, the fact that percussion sounds consist, to a good approximation, of damped oscillations which are easy to produce electronically, makes it worthwhile to generate these sounds separately, by means of dedicated hardware. This solution is intended to reduce the workload of the voice modules, so that they can be more efficiently allocated to the synthesis of sounds of higher complexity. It also offers the benefit of simpler control, as will appear below.

Figure 25 shows the basic percussion circuit configuration, which consists of a trigger/envelope driver and a phase-shift oscillator. Normally switch Q1 is open and C1 is charged at -15 V. The base network of Q2 is such that, under these conditions, the circuit is unable to sustain oscillation due to insufficient base bias to keep Q2 on.

The arrival of a trigger pulse causes Q1 to close and discharge C1 to ground. The voltage step now appearing across R2 is large enough to bias Q2 in the active region, where the circuit is capable of sustained oscillations. The frequency of oscillation is determined by the components included in the feedback network of Q2.

After removal of the trigger pulse, Q1 is opened and C1 is allowed to charge again to -15 V, at a rate determined by R1, R2 and C1 itself. This voltage transition gradually turns Q2 off, via R2, and causes the amplitude of oscillation to undergo a smooth decay to zero. The rate of decay is determined by the rate at which C1 charges to -15 V. Resistors R3 and R4

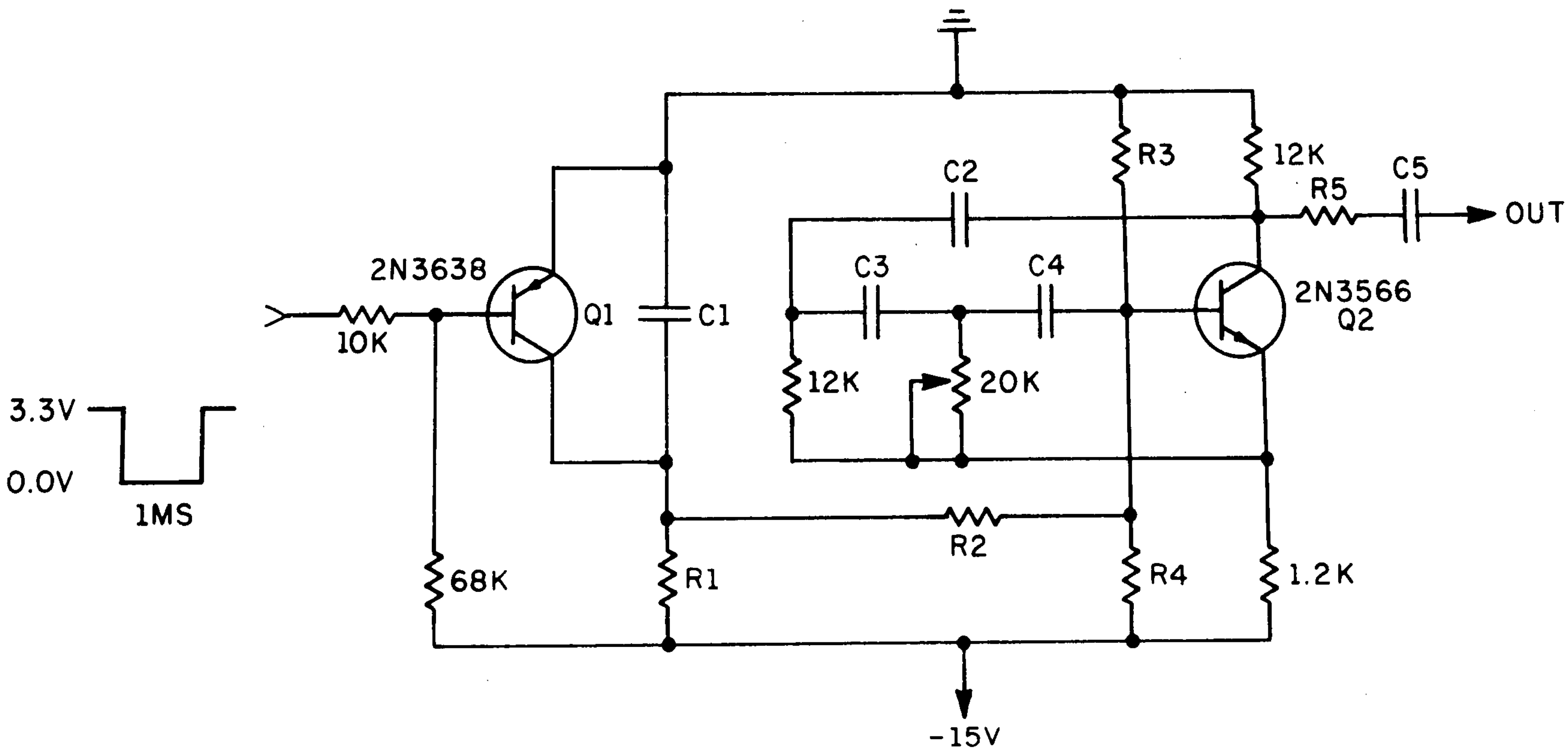


Figure 25. Basic Percussion Circuit.

Figure 25. Basic Percussion Circuit.

set the point at which the amplitude of oscillation reaches the zero value. The trimmer shown is used for final tuning of the percussion pitch.

Thus, whenever the circuit is struck by a trigger pulse, it emits a series of damped oscillations. By properly choosing the various component values, the same basic configuration can be easily adapted to span an entire family of percussion instruments like bass drums, tom-toms, bongos, claves, wood-blocks, congas and castanets. As an example, the following are the component values that have been used to synthesize a bongo:

R1 = 220 k	C1 = .05 μF
R2 = 1 M	C2 = .0033 μF
R3 = 2.2 M	C3 = .0033 μF
R4 = 100 k	C4 = .0033 μF
R5 = 150 k	C5 = 750 pF

Figure 26 shows the simplified circuit schematic of an entire percussion ensemble. The control portion involves two read/write memories addressed in parallel by a variable-modulus counter. The information residing in the lower memory specifies the pattern according to which the instruments are to be struck in sequence. The information of the upper memory determines the intensity with which each instrument is to be struck.

Circuit operation is rather straightforward. The arrival of a clock pulse from the orchestra duration control steps the counter and causes two new words of data to appear at the outputs of the respective memories. The word from the upper memory, after D/A conversion, programs the attenuator for the proper loudness. The word from the lower memory sets up the decoder so that the 1 ms trigger pulse emitted by the one-shot is directed to the proper instrument. Thus a sequence of clock pulses results in the generation of a pattern of percussion sounds. Pattern length is determined by the modulus of the counter, which can be programmed externally.

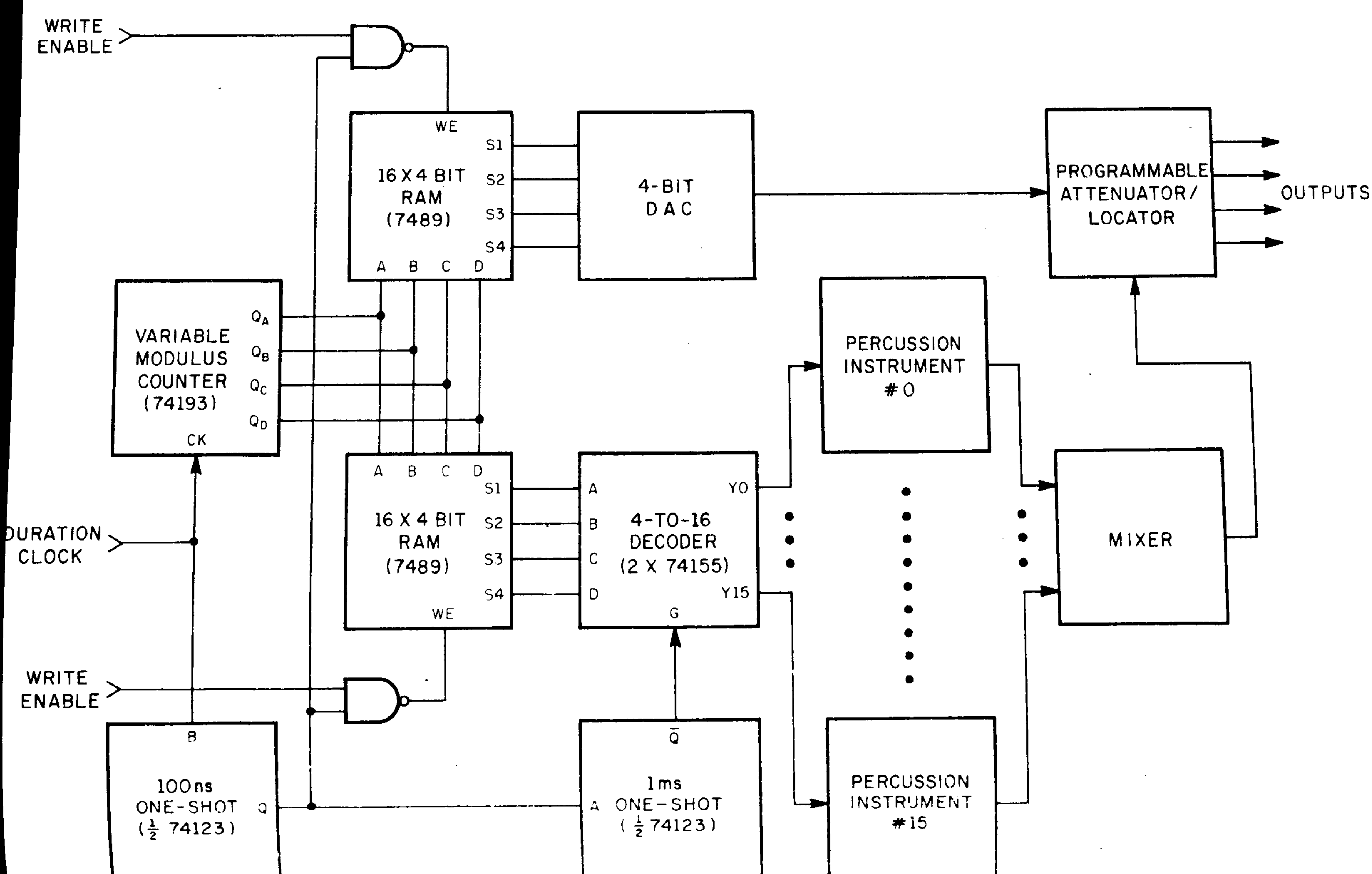


Figure 26. Percussion Ensemble.

Being provided with read/write capabilities, both memories can be independently updated to accommodate new patterns. This fact, together with the programmability of the counter, makes the circuit suitable for generating a wide variety of rhythms and patterns. Memory updating can be carried out while the percussion ensemble is in action. To avoid synchronization problems, the leading edge of the duration clock is followed by the emission of a sharp pulse by the 100 ns one-shot, which strobes the information waiting at the memory data-inputs into the memory register being currently addressed. The entire percussion ensemble is silenced by disabling the duration clock.

Figure 26. Percussion Ensemble.

VI. SOUND DISTRIBUTION AND SWITCHING

6.1 Spatial Sound Distribution.

The four-channel output system discussed earlier is primarily intended for the control of sound location in studio-type situations. For operation in concert-hall environments, the output capabilities of the instrument are expanded by the addition of 24 discrete channels, each complete with power amplifying hardware to drive a corresponding loudspeaker. Spatial elements are then controlled by routing sounds to the various loudspeakers which are placed at different locations throughout the performance area.

This output configuration can handle spatial situations of much greater complexity than the quadraphonic system. One example is offered by sound traffic, an effect that can be readily produced by switching a sound from one position to the next, along a chain of consecutive loudspeakers. By permuting the order in which the loudspeakers are grouped into chains, one can specify arbitrary paths of sound traffic.

Sound routing is handled by means of analog gates, or audio gates, as they will more properly be termed henceforth. Being a two-state device, a gate requires only one bit of control information, and this is a highly desirable feature in that it maintains the overall information requirements within manageable limits. Thus, a complete sound routing system necessitates 24 audio gates, each gate being assigned to control the presence or absence of sound on a corresponding channel. Accordingly, the specification of an entire sound distribution pattern can be done with 24 bits of data.

To set a distribution pattern in motion, the associated information must be properly processed, and a suitable device to perform this task is a circular shift-register. The register consists of 24 cells, each cell being assigned to control the state of a corresponding audio gate. It is also

provided with parallel load facilities for the direct entry of sound distribution patterns. The order in which the gates are assigned to consecutive register cells defines a corresponding chain of loudspeakers extending throughout the performance space. Thus, shifting a bit pattern around the register causes an identical sound pattern to move around the loudspeaker chain. The direction of motion is determined by the direction of shift.

To offer a wider choice of sound traffic paths, the shift-register is broken down into six shorter registers, as shown in Figure 27. More complex sound paths are then specified by permuting the order in which the shorter registers are joined together. Permutation patterns are programmed by means of a 6 x 6 interconnection matrix. When the (i,j)-th bit of the matrix is at logic one, the serial output from the i-th register is diverted to the serial input of the j-th register. Thus each register can be connected to any other register, including itself.

Although this scheme does not exhaust all possible traffic paths that can be specified with a system of 24 loudspeakers, it offers enough variety to satisfy a wide range of musical needs without rendering control unclear or too complex.

To preserve its spatial individuality, each of the four orchestras is provided with an independent sound distribution system, so that it can control the location and movement of its sounds according to its own patterns and rates.* A block diagram of the entire sound distribution system is shown in Figure 28. The system has four input busses, each of which carries the sounds pertaining to one of the four orchestras. These busses are driven by the quadraphonic output system discussed earlier. The assignment of each orchestra to one of the quadraphonic outputs is accomplished by proper

* The sound distribution control system was designed by J. Divilbiss and R. Borovec.

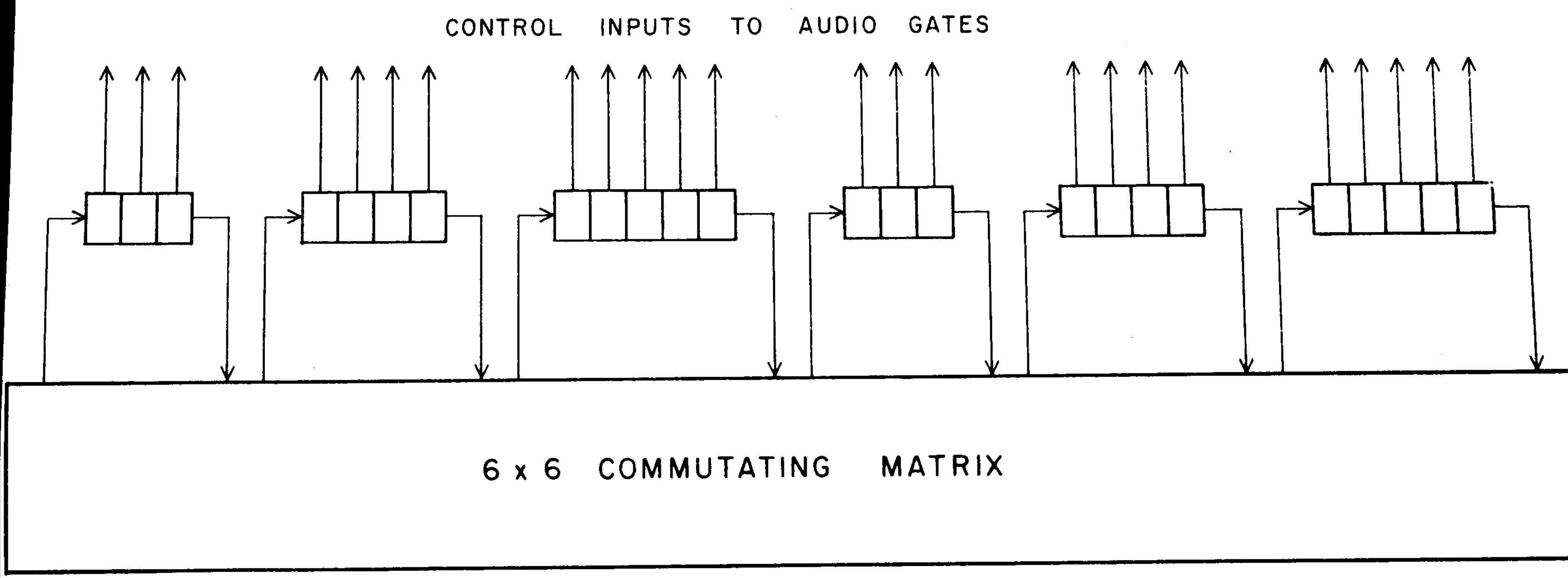


Figure 27. Distribution Control System.

Figure 27. Distribution Control System.

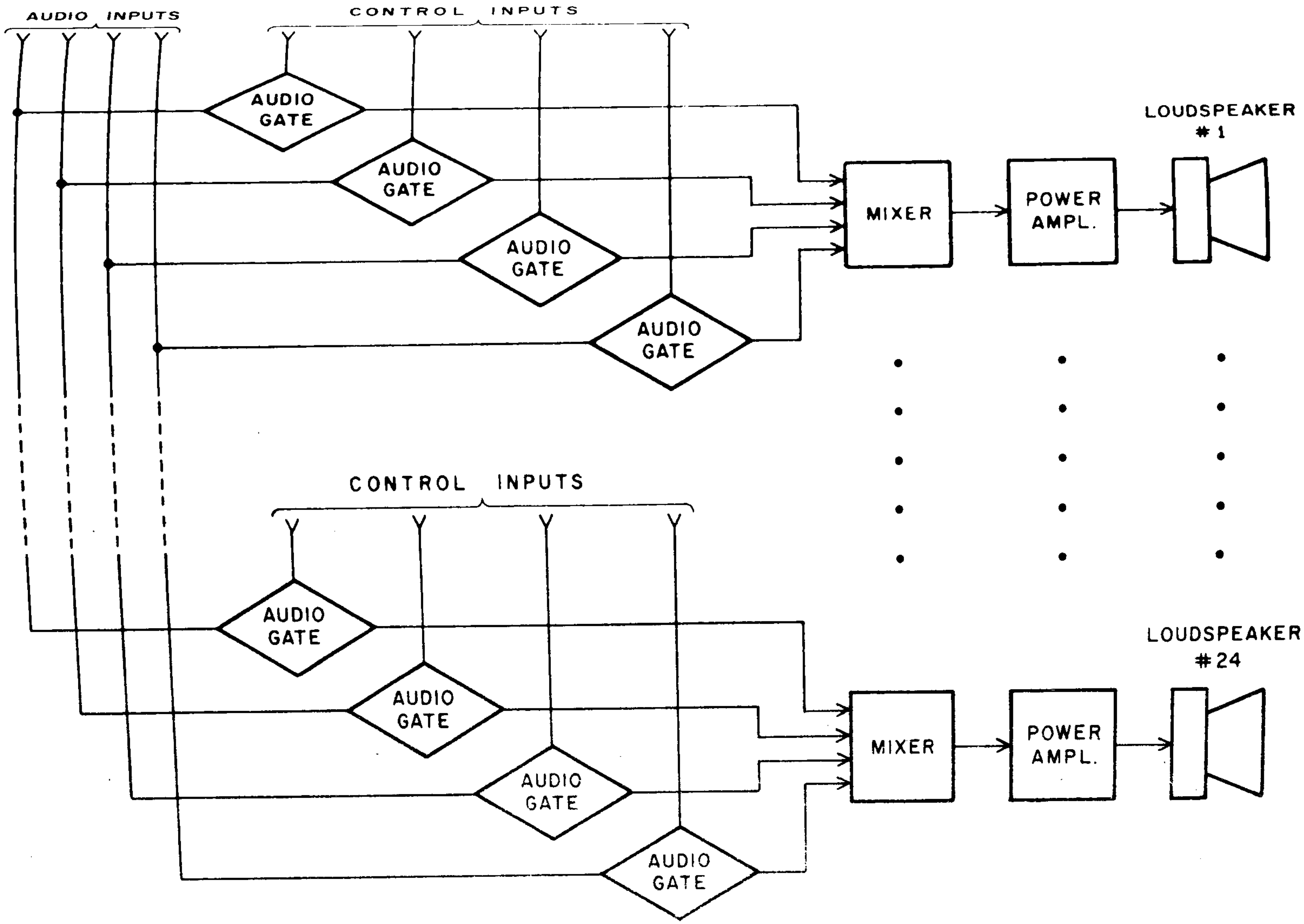


Figure 28. Sound Distribution System.

initialization of the directional control inputs to the various attenuator/ locator devices.

6.2 Audio Switching.

As shown in the block diagram of Figure 28, the realization of the sound distribution system requires a total of 96 audio gates. With a number of this magnitude, the cost per gate plays an important role in the choice of the circuit realization to be adopted.

An audio gate, besides satisfying such obvious requirements as low distortion, high on-to-off transfer ratio, and low control signal feedthrough, must also exhibit specific transient characteristics in order to ensure proper audio switching. The last requirement is motivated by the fact that when a sound is switched on or off, spurious partials are generated which may alter the tonal character of the sound considerably, as has been discussed in connection with sound enveloping.

As Fourier analysis reveals, the amount of spurious partials accompanying sound switching usually increases with the rate at which sound builds up or decays. Thus, unwanted switching effects can be easily reduced by employing gates with low switching speeds. From the viewpoint of sound movement control, however, it is desirable to have fast gates so that sounds can be switched around the performance space at arbitrary rates. A compromise between the two conflicting requirements can be determined experimentally.

Analog gates based on solid-state, electro-optical devices like Raysistors or Vactrols usually exhibit predetermined and highly asymmetrical switching characteristics which render these components unsuited to the present application. The cost of these devices at the time of design was also a key factor that contributed to their rejection. Analog gates of the FET or MOS

type must be ruled out because of their high switching speeds and the lack of a simple and inexpensive circuit configuration to slow them down.

The specifications set forward above have been met with variable-transconductance transistor pairs. The switching characteristics of these devices are set to their prescribed values by means of simple RC networks, as shown in Figure 29. To illustrate circuit operation, consider the first gate, which consists of transistor pair Q1/Q2 and biasing transistor Q4. The gain or transmission of the gate is determined by the collector current of Q4. Since Q4 is connected as an emitter follower, its collector current and, hence, the transmission of the gate, is modulated by the voltage applied to the base of Q4. This voltage is in turn controlled by current switch Q3. When the control input is high, Q3 is on and its collector current develops a voltage drop across D2 and R2 in series, which biases Q4 on also. When the control input is low, transistor Q3 is off, and this causes Q4 to be off, due to the absence of collector current from Q3 to keep the base network of Q4 biased.

Although Q3 is switched from one state to the other at logic speed, the voltage transitions at the base of Q4 are slowed down by damping capacitor C. Thus, gate transmission builds up and decays according to the transient characteristics of an RC element. The switching time-constant equals $(R2 + r_D) \cdot C \doteq 12$ ms, where r_D is the dynamic resistance of diode D1. This time-constant value represents a satisfactory compromise between the need for fast sound movement and that for spurious partials reduction.

Diode D1 compensates for the presence of the base-to-emitter voltage drop of Q4. Hence it prevents the collector current of Q4 from falling too rapidly to zero when the voltage at the base of Q4 decays below the value of

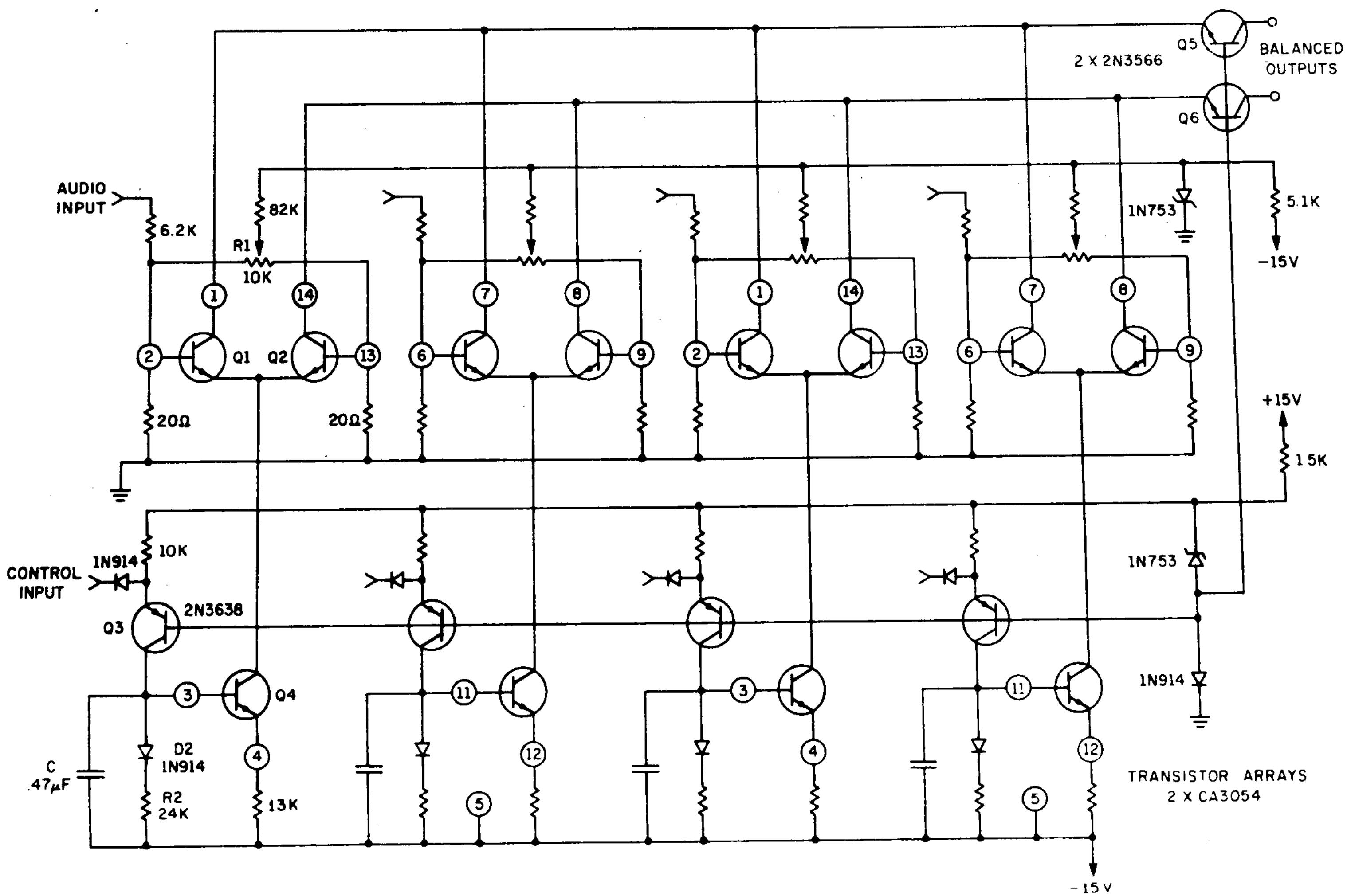


Figure 29. Quadruple Audio Gate.

a junction voltage drop. Trimmer R1 is used to adjust for zero dc offset at the gate output.

The outputs from the four gates are mixed together by tying corresponding collectors of the transistor pairs in parallel. The circuit realization of a quadruple audio gate requires two CA3054 transistor arrays, as shown.

The balanced outputs from the gates are buffered to the subsequent stage by common-base transistors Q5 and Q6. The purpose of these transistors is to keep the collectors of the gates near zero potential and, therefore, minimize the power dissipated by the transconductance pairs. This precaution is taken to reduce temperature variations of the dc output offset which constitutes the primary cause of control signal feedthrough. Indeed, when a gate is switched on or off, its dc offset is switched on or off as well, resulting in a form of coupling of the control input to the output. Control feedthrough is particularly noticeable when the distribution controller is programmed for fast sound movement and loudness control is, at the same time, programmed for a rest. This problem does not arise in connection with the programmable attenuator/locator because, whenever the attenuator is programmed for a rest, the transconductance pairs of the locator are automatically off. With the precaution illustrated above, control signal feedthrough is maintained more than 60 db below full signal level.

The task of converting the balanced outputs from the buffers into a single-ended signal suitable to drive an 8Ω-loudspeaker is accomplished by the circuit of Figure 30. The circuit consists of the usual differential amplifier and a power stage to boost the amplifier drive capability. The trimmer shown is used to adjust the quiescent current of the power stage for the prevention of thermal runaway. Power stage and loudspeaker are dc

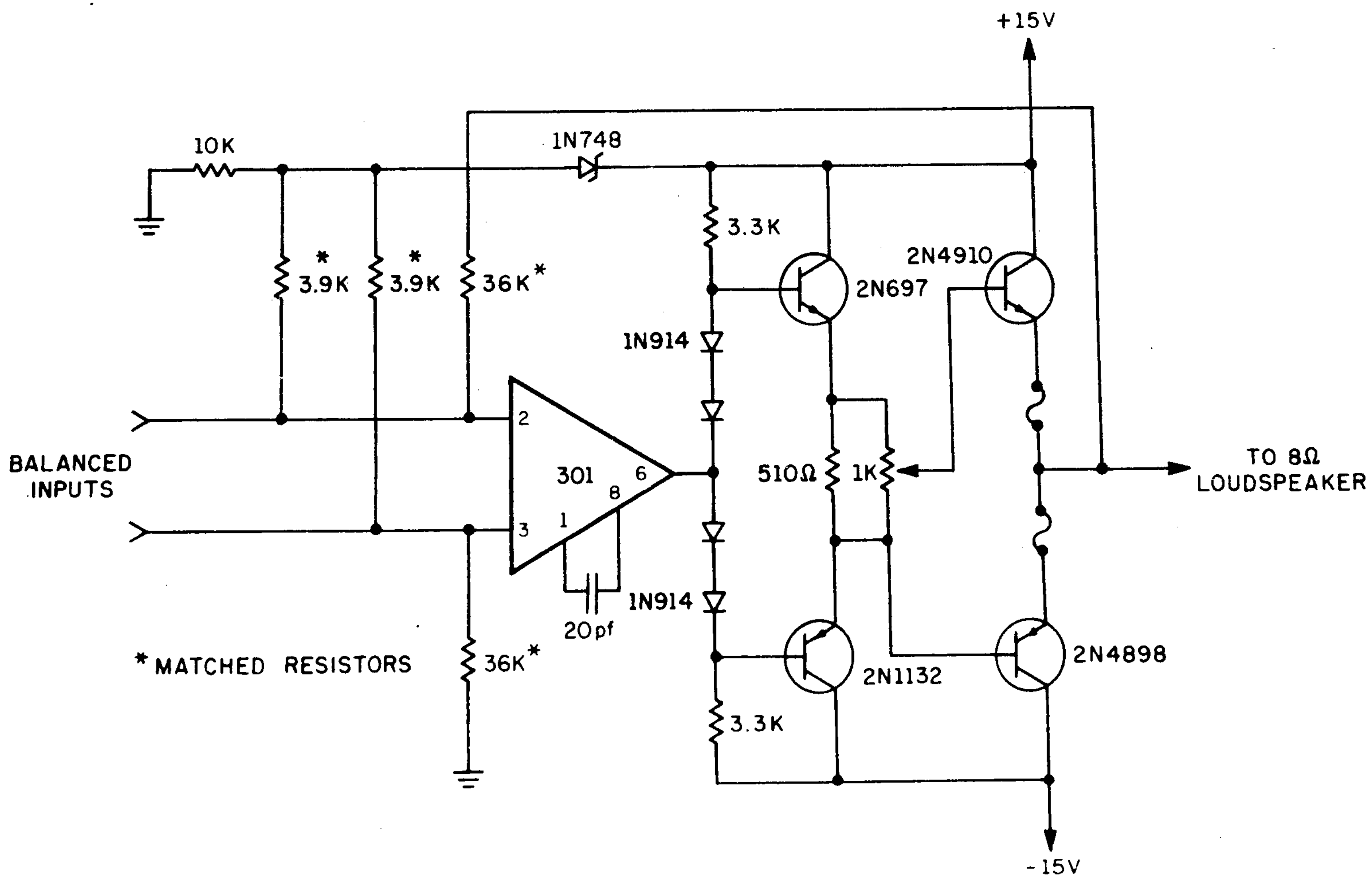


Figure 30. Power Amplifier.

coupled to prevent the use of a bulky decoupling capacitor. Two 1A-fuses are used to protect the loudspeaker against amplifier malfunction.

The circuits of Figures 29 and 30 are easily accommodated on a single printed-circuit board of standard size (4 1/2" x 6").

Figure 30. Power Amplifier.

VII. CONTROL SYSTEM

7.1 Musical Parameter Control.

The material presented so far has dealt primarily with circuits and techniques for the synthesis and processing of sounds. To make musical parameter control possible, all modules encountered respond to some form of digital or analog control which allows their electronic characteristics to be programmed externally. We shall now examine in greater detail how these circuit parameters are actually programmed in order to produce music.

As stated at the beginning of this work, all control functions are specified in purely digital form. Several of the sound modules encountered, namely, equitempered pitch generators, duration controllers, envelope generators and digital waveshape generators, are already equipped with digitally programmable inputs. All that needs to be done, then, is to feed these control inputs with a stream of binary information of the proper word length. Most of the parameters of musical interest, however, are of a continuous, rather than a discrete nature, as reflected by the fact that the corresponding circuit characteristics are controlled by analog voltages instead of digital signals. To make these parameters accessible to digital control, analog sound modules must be preceded by a proper digital-to-analog interface. This task is accomplished by means of piece-wise linear function generators of the same type as those used in the generation of envelopes. Although these devices were discussed in connection with loudness control because that is probably their most illustrative application, piece-wise function generators are of far more general applicability, in that they can be used to control the time evolution of any analog parameter.

This generalized notion of envelope for the uniform control of all continuous quantities results in a high degree of control as well as hardware

standardization. The use of this type of interface to control analog modules also results in a significant data-rate reduction, since the parameters to be controlled are the envelope breakpoints. Indeed, had the interface consisted of straight D/A converters, the quantities to be specified would have been individual envelope samples, a task that involves a much higher data-rate than the mere specification of breakpoints.

In the light of the above discussion, all sound modules may now be treated as programmable devices that respond to purely digital control. Regardless of their nature, significance and word-length, the dynamic control of musical parameters is accomplished in a uniform manner by means of binary sequences. As implied by its name, a control sequence consists of a series of binary words which are successively spaced in time at determinate, but not necessarily uniform, time intervals. This control approach is intended to be of enough generality as to leave the programming of sound modules open to a broad choice of digital systems and techniques for its implementation.

Binary sequences are generated by programmable word generators and by the human operator. Depending on the structural level at which they affect music, control sequences are classified as microstructural or macrostructural.

Microstructural sequences are those effecting individual parameter changes and, hence, are the sequences that control the characteristics of sound modules directly, as shown in Figure 31. Since perceptual parameters change at rates in the range of 1 to 100 Hz, these sequences are too fast for human control in real-time. Hence microstructural sequences are generated by word generators that are built into the system.

Macrostructural sequences are used to manipulate microstructural sequences for the creation of more complex musical structures. To make this

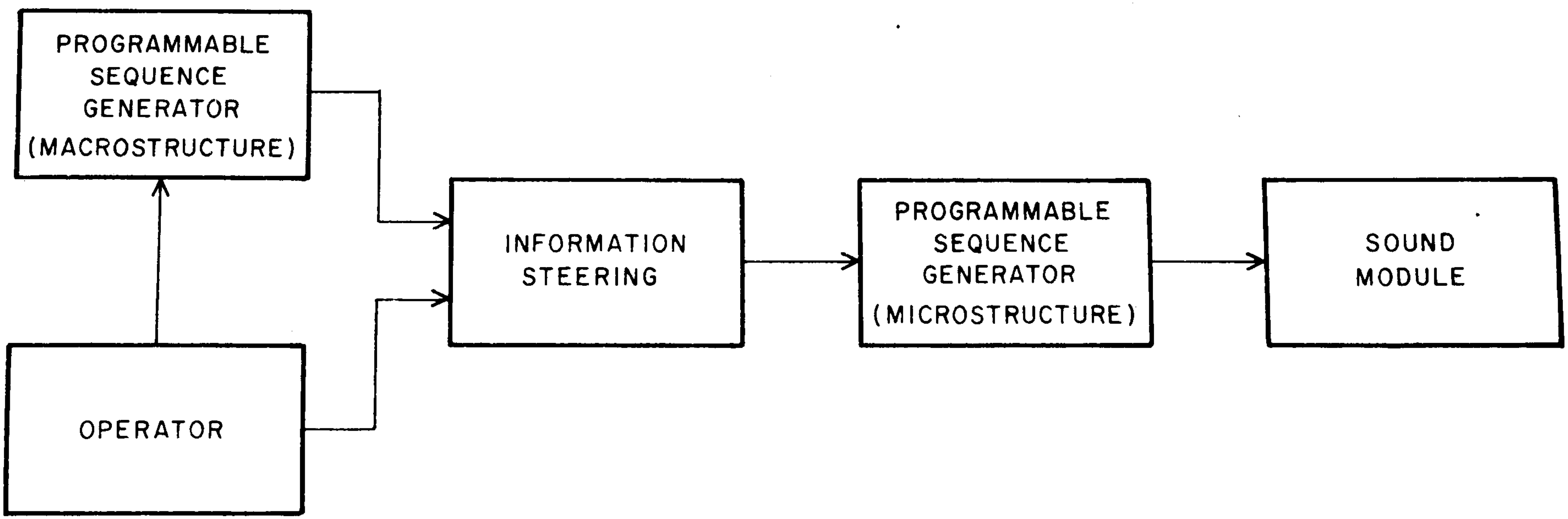


Figure 31. Block Diagram of Sound Module Control.

manipulation possible, all microstructural sequence generators are themselves externally programmable. Since they affect music at a higher level of structural complexity, macrostructural sequences are characterized by lower timing rates, which may typically be of the order of .01 to 1 Hz. These are rates at which a human operator can comfortably perform control actions in real-time. Hence, the generation of macrostructural sequences for the control of microstructural sequence generators is made accessible to the operator.

If so desired, the operator can delegate the generation of macrostructural sequences to the system itself. To handle this task, the latter is also provided with macrostructural sequence generators, in addition to the microstructural ones already discussed. The information associated with these macrostructural sequences is called automatic, as opposed to that generated by the operator which is called manual. The specification of whether macrostructural information is to be generated manually or automatically is expressed by the operator himself in terms of one-bit sequences to control the information steering device, which consists of data selectors. To allow higher levels of control complexity, macrostructural sequence generators are also programmable. However, their programming is done by the operator exclusively.

7.2 Programmable Binary Sequence Generation.

A device that has proved to be particularly attractive for the generation of binary control sequences, both in terms of simplicity and cost, is the feedback shift-register. As shown in Figure 32 for the case of 4-bit word sequences, this device consists of a shift-register and a combinational circuit which computes the next value of the serial input as a function of

Figure 31. Block Diagram of Sound Module Control.

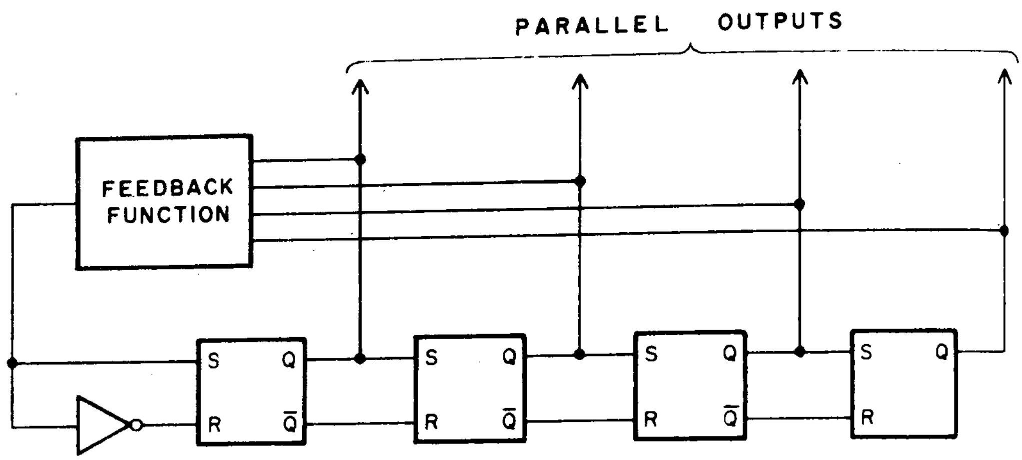


Figure 32. Feedback Shift-register.

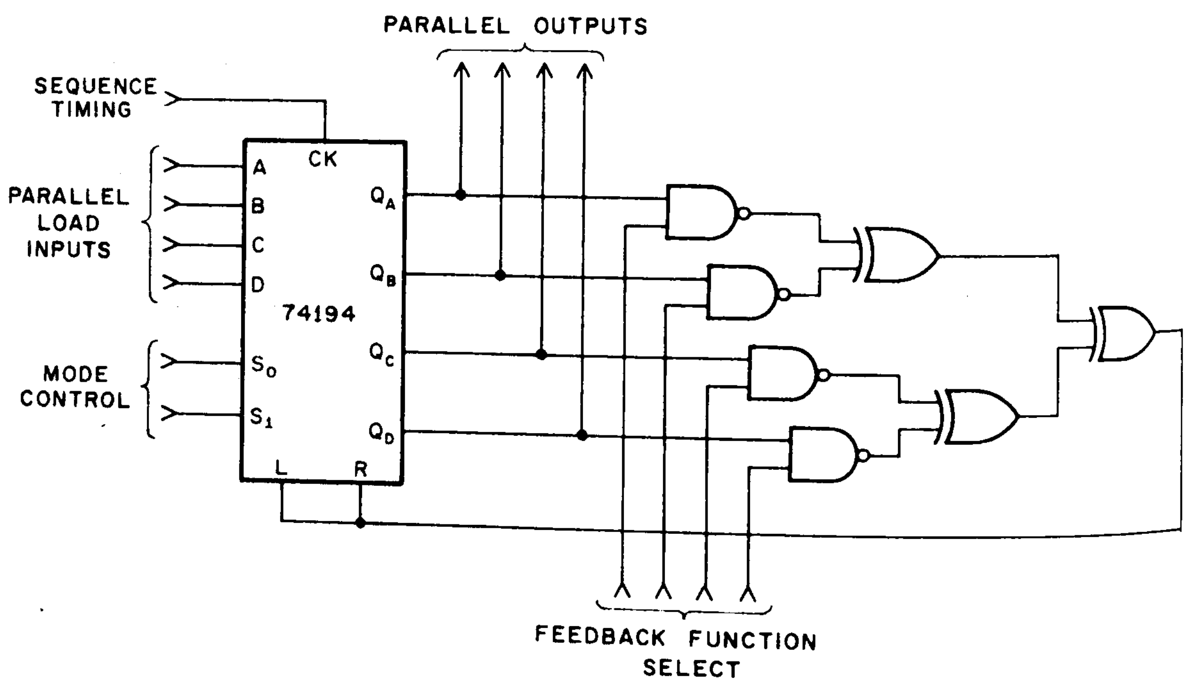


Figure 33. Programmable Sequence Generator.

the present state of the register. At every shift-pulse the register undergoes a transition from one state to another, so that a series of shifts carries the register through an entire sequence of states. Thus, by using a shift-register with parallel output facilities, the sequence of words appearing at the output lines can be used to control musical parameters directly.

Shift-register sequences have been widely investigated in the literature[34] and will not be discussed in detail here. It may suffice to point out that, in general, with a given feedback function, a register can generate a whole family of different sequences, and that different feedback functions yield different families of sequences. Thus, the versatility of a feedback shift-register can be considerably expanded if a programmable rather than a fixed feedback function is used.

Once a particular feedback function has been programmed, the register can be initialized to any one of the sequences of the corresponding family by loading it with a word of that sequence. To make this possible, the register is provided with parallel-load facilities. The use of a bidirectional shift-register also increases the versatility of this device, since the sequences resulting from left shifts are generally different from those obtained with right shifts.

Feedback shift-register sequences enjoy the musically attractive feature of being periodic. This means that after a sufficient number of shifts, a sequence repeats itself over and over again, until the system is programmed for a different sequence.

A convenient way of implementing a programmable feedback function is by means of a modulo-2 adder with programmable patterns of active inputs. It can be easily shown[27] that the periodic lengths of the sequences obtained

with this class of feedback functions are bounded by 2^n-1, where n is the shift-register length. Figure 33 shows the realization of a 4-bit word sequence generator based on the above ideas.

By exploiting the options made available by the various control inputs, the same device can be programmed to generate a broad class of control sequences. In general, each sequence posesses a musical significance of its own. Hence, the generator offers a set of basic compositional elements which can be easily manipulated by the improviser to create more complex musical structures.

In order to satisfy a class of musical needs as broad as possible, it is necessary that musical parameters be programmable over wide dynamic ranges. However, in defining a particular musical context, the composer is likely to use only limited portions of the dynamic ranges available. Thus, control sequences can be musically useful only if the associated stream of information can be constrained within appropriate limits or windows.

Since feedback shift-registers do not offer facilities for information biasing as such, this task must be handled independently. A simple way to achieve this goal is to partition the bits of a control word into a most-significant and a least-significant group, and to handle the two groups separately, the first with a macrostructural sequence and the second with a microstructural sequence. This causes the resulting information to be biased within windows whose width is determined by the length of the least-significant portion of each word, and whose position within the total range is controlled by the most-significant data.

Although adequate for the control of certain parameters, this scheme is not sophisticated enough for others, due to the fact that windows are constrained to be of fixed widths.

A more flexible scheme of information biasing relies on programmable window detectors, as shown in Figure 34 for the case of 4-bit words. The detector output results in an assertion only if the input information falls within the limits prescribed at the control inputs. Hence, this device can readily be incorporated into an information-filtering scheme for the biasing of entire control sequences within arbitrarily programmable windows.

Since sequence generators that are equipped with all the programming facilities illustrated above possess a substantial number of control inputs, only a centralized set of devices of this type is used. The related sequences are then transferred to local memories, where they are recirculated for dynamic parameter control. These peripheral memories consist of read/write IC memories addressed by programmable-modulus counters, in the manner illustrated in connection with the control of percussion sounds.

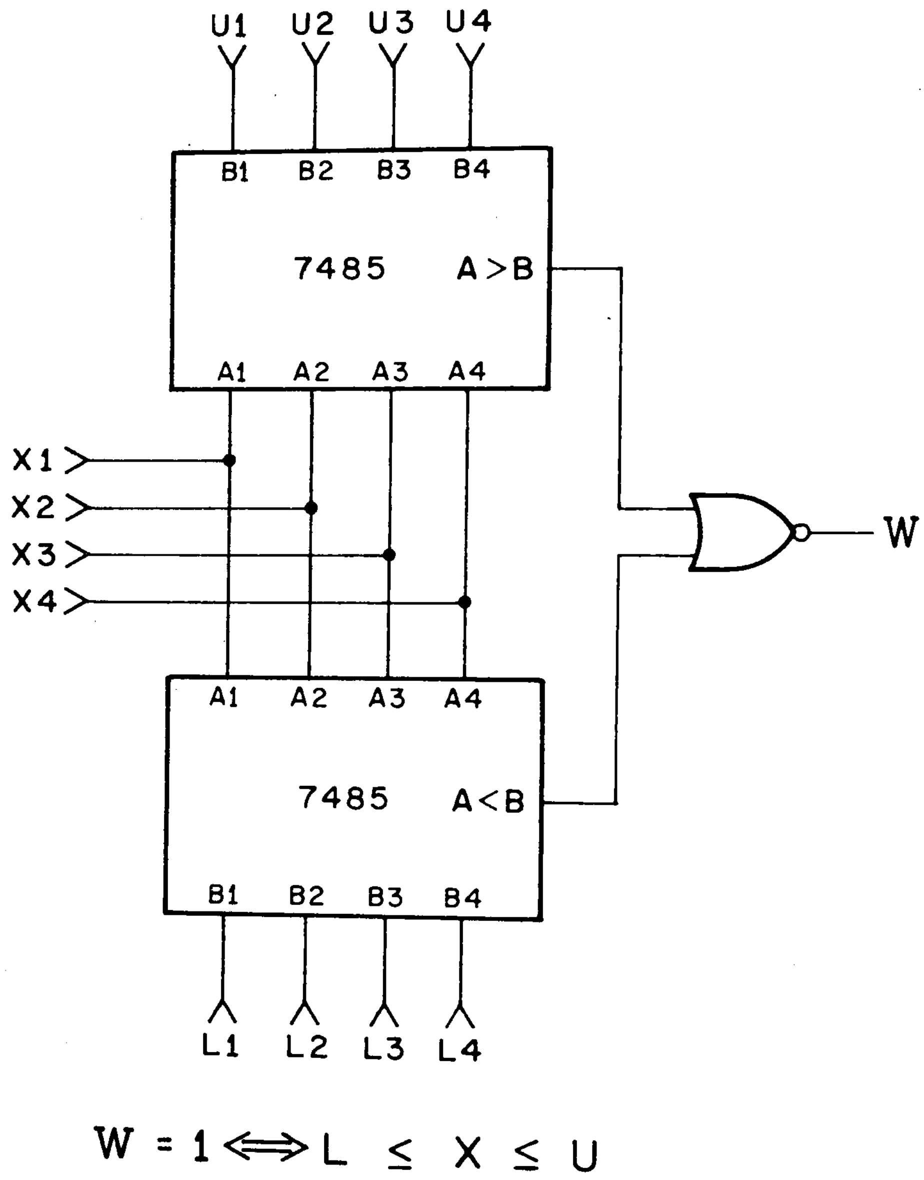


Figure 34. Digital Window Detector.

VIII. CONCLUSION

A hardware system has been presented which allows for the composition of electronic music in real-time. Real-time operation is achieved by a proper functional partition which assigns analog hardware to the synthesis and processing of sounds and digital hardware to the control of musical parameters. Throughout the design of the system particular emphasis has been placed on assuring wide ranges of sound quality as well as flexibility of parameter control.

The use of sophisticated hardware made available by recent technology has led to the development of sound modules that afford a much richer range of tonal and instrumental possibilities than obtainable from present-day analog synthesizers. Furthermore, the output configuration based on a system of loudspeakers distributed throughout the performance space offers a relatively new approach to the spatial characterization of sounds and ensures a total immersion and involvement of the listener with the sound/space environment.

The adoption of generalized envelope functions for the dynamic control of continuous musical quantities introduces a high degree of control uniformity, in that all parameters, be they discrete or continuous, can be handled in the same manner. Parameter control is based on the notion of discrete control sequences, which are characterized by clocking rates of about 100 Hz or less. While involving a relatively low data-rate, this approach is also of enough generality as not to put definite constraints on the type of digital systems to be used for the control of sound modules.

The structural nature of music allows a correspondingly structured organization of control which is particularly significant in the real-time

interaction between composer/performer and instrument. In this respect, a general distinction between microstructural and macrostructural control sequences has been emphasized so that a proper balance between operator responsibilities and machine responsibilities can be achieved.

The fact that technology continues to market analog and digital systems of increasing complexity and competitive cost is bound to have a significant effect on electronic music. For instance, with the recent appearance of microcomputers, it is now becoming economically feasible to use one or more of these devices to program sound modules, so that the sophistication of the control system outlined in this work can be expanded further.[28] Technology advances are bound to affect also sound processing hardware and techniques. In this respect, it is highly desirable that sound modules become more digitally oriented, since digital systems enjoy the intrinsic property of being far more accurate and stable than analog circuits. In the light of these remarks, it seems appropriate to speculate that the electronic music systems of the next generation will be based on a central, general-purpose computer controlling a collection of peripheral, special-purpose computers whose task is the synthesis and processing of sounds.

LIST OF REFERENCES

1. Moore, F. R., "Computer Controlled Analog Synthesizers," Bell Telephone Laboratories, Inc. (Internal Memo.).

2. Moog, R. A., "Voltage Controlled Electronic Music Modules," J. Audio Engineering Society, Vol. 13, p. 200, 1965).

3. Moog, R. A., "Electronic Music: Its Composition and Performance," Electronics World, February 1967.

4. Friend, D., "The ARP Synthesizer: A New Instrument for Musical Composition and Performance," AES Preprint #742 (B-3), Audio Engineering Society 39th Convention, October 12-15, 1970.

5. Mayer, A., "The PUTNEY: A New Generation of Synthesizers," AES Preprint #750 (B-1), Audio Engineering Society 39th Convention, October 12-15, 1970.

6. Mathews, M. V., <u>The Technology of Computer Music</u> (MIT Press, Cambridge, Mass. 1969).

7. Mathews, M. V., F. R. Moore and J. C. Risset, "Technology and Future Music," Bell Telephone Laboratories, Inc. (Internal Memo.)

8. Beauchamp, J. W., "Electronic Music Equipment: 1950-1970," Experimental Music Studio Technical Report No. 21, University of Illinois, February 1970.

9. Risset, J. C. and M. V. Mathews, "Analysis of Musical Instrument Tones," Physics Today, Vol. 22, No. 2, pp. 23-30, 1969.

10. Chowning, J. M., "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation," J. Audio Engineering Society, Vol. 21, pp. 526-534, September 1973.

11. Gibbons, J. F. and H. S. Horn, "A Circuit with Logarithmic Transfer Response over 9 Decades," IEEE Trans. Circuit Theory, Vol. CT-11, pp. 378-384, September 1964.

12. Borlase, W. and E. David, "Log Circuits," Analog Devices, Inc., Application Note, March 1969.

13. Gilbert, B., "A Precise Four Quadrant Multiplier with Subnanosecond Response," IEEE J. Solid-State Circuits, Vol. SC-3, pp. 365-373, December 1968.

14. Gilbert, B., "Comment on 'Single-Ended-Input Single-Ended-Output Four-Quadrant Analog Multiplier'," IEEE J. Solid-State Circuits, Vol. SC-7, p. 434, October 1972.

15. Mitra, S. K. (ed.), "Active Inductorless Filter," IEEE Press, Inc., New York, 1971.

16. Capella, D., "FETs as Voltage-Controlled Resistors," Siliconix, Inc., Application Note, February 1973.

17. Bruton, L. T., "Electronically Tunable Analog Active Filters," IEEE Trans. Circuit Theory, Vol. CT-19, pp. 299-301, May 1972.

18. Sparkes, R. G. and A. S. Sedra, "Programmable Active Filters," IEEE J. Solid-State Circuits, Vol. SC-8, pp. 93-95, February 1973.

19. Kaehler, J. A., "Periodic-Switched Filter Networks - A Means of Amplifying and Varying Transfer Functions," IEEE J. Solid-State Circuits, Vol. SC-4, pp. 225-230, August 1969.

20. Bruton, L. T. and R. T. Pederson, "Electronically Tunable Time-Multiplexed Active Filters," IEEE J. Solid-State Circuits, Vol. SC-7, pp. 266-267, June 1972.

21. Kerwin, W. J., L. P. Huelsman, and R. W. Newcomb, "State-Variable Synthesis for Insensitive Integrated Circuit Transfer Functions," IEEE J. Solid-State Circuits, Vol. SC-2, pp. 87-92, September 1967.

22. Thomas, L. C., "The Biquad: Part I - Some Practical Design Considerations," IEEE Trans. Circuit Theory, Vol. CT-18, pp. 350-457, May 1971.

23. Geffe, P. R., "Toward High Stability in Active Filters," IEEE Spectrum, Vol. 7, pp. 63-66, May 1970.

24. Oneil, B., "A Precision Four Quadrant Multiplier," Intersil, Inc., Application Bulletin A011, June 1972.

25. Wilson, G. R., "A Monolithic Junction FET-NPN Operational Amplifier," 1968 Internat'l. Solid-State Circuits Conf., Digest of Tech. Papers, pp. 20-21.

26. Wittlinger, H. A., "Applications of the CA3080 High-Performance Operational Transconductance Amplifiers," RCA's Linear Integrated Circuits Application Note ICAN-6668, March 1972.

27. Golomb, S. W., <u>Shift Register Sequences</u>, Holden-Day, Inc., San Francisco, 1967.

28. Franco, S., "The Oberlin Hybrid Project," Oberlin Conservatory of Music Internal Memo, June 1974.

VITA

Sergio Franco was born in Capriva del Friuli, a rural village of northeastern Italy. After graduation in Physics from the University of Rome, he taught high school Mathematics and Physics in Rome. In 1967 he was awarded a Fulbright Fellowship to continue his studies in the USA. He first joined the Department of Physics of Clark University, Worcester, Massachusetts, as a teaching assistant, where he earned a M.A. degree in Physics in the area of the Mössbauer effect. In 1968 and 1969 he was associated with the Illiac III Project at the Department of Computer Science of the University of Illinois at Urbana, where he designed the analog circuitry of the Scan/Display System. After attending graduate courses in Computer Science at the University of Toronto during the years 1970 and 1971, he returned to the Urbana campus, where he became associated with the Coordinated Science Laboratory and the School of Music. Since then, his main interest has been in the area of electronic music. During 1971 and 1972 he worked with the composer, Salvatore Martirano, in the development of the musical system described in this thesis. In 1973 he joined the Oberlin Conservatory of Music where he teaches courses in analog and digital design of Electronic Music Systems, and is currently developing the Oberlin Hybrid Project, a computer controlled system for the composition of music in real-time. Mr. Franco has published a number of articles in Physics and in Electronics.

BIBLIOGRAPHIC DATA SHEET

1. Report No. UIUCDCS-R-74-677	2.	3. Recipient's Accession No.
4. Title and Subtitle HARDWARE DESIGN OF A REAL-TIME MUSICAL SYSTEM		5. Report Date December 1974
		6.
7. Author(s) Sergio Franco		8. Performing Organization Rept. No. UIUCDCS-R-74-677
9. Performing Organization Name and Address Department of Computer Science University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. Project/Task/Work Unit No.
		11. Contract/Grant No.
12. Sponsoring Organization Name and Address		13. Type of Report & Period Covered technical
		14.

15. Supplementary Notes

16. Abstracts

This thesis describes the hardware design of a hybrid system for the composition and performance of electronic music in real-time. While analog circuitry is primarily employed in the generation and processing of sounds, digital circuitry is devoted to the exercise of control, under the immediate supervision of the composer/performer. Such a digital vs. analog partition, together with a proper choice of the man vs. machine interface, is intended to satisfy the much emphasized musical need for the immediate, real-time interaction between the composer/performer and his instrument.

17. Key Words and Document Analysis. 17a. Descriptors

Electronic Music
Real-Time Music Synthesis
Musical Waveshape Systhesis/Control

17b. Identifiers/Open-Ended Terms

17c. COSATI Field/Group

18. Availability Statement

Unlimited

19. Security Class (This

21. No. of Pages